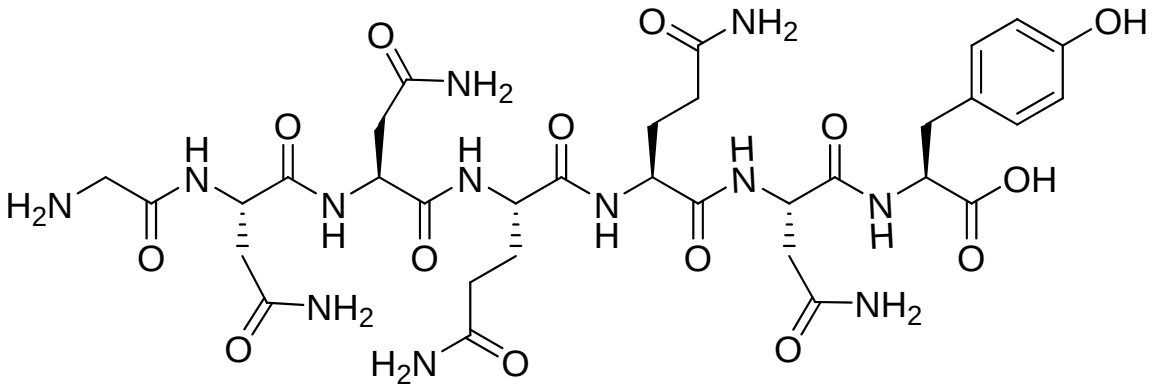
# The Driving Forces of Peptide Aggregation: A Study of the Yeast Sup35 Prion Fragment GNNQQNY

Author: Kevin Lebo

# The Driving Forces of Peptide Aggregation:
## A Study of the Yeast Sup35 Prion Fragment GNNQQNY



Kevin Lebo

Advisor: Dr. Jianmin Gao
Boston College Chemistry Department

2008

## Table of Contents

## Abstract:

Protein aggregation can be highly detrimental to organisms, and has been associated with diseases including Alzheimer's, Huntington's, type II diabetes, and transmissible spongiform encephalopathies such as Mad Cow disease. There is no single amino acid sequence responsible for aggregation into amyloid-like structures, but rather a large range of amyloidogenic peptides have been discovered. A fragment of the yeast Sup35 prion, GNNQQNY, has been found to aggregate using a "dry, steric zipper" structure. This study looks at mutants of GNNQQNY in order to elucidate the exact contributions of various amino acids to the aggregation process. The data shows that Gly1 is important to the overall aggregation of the peptide, and that a deletion of the amino acid reduces the formation of larger amyloid structures. Mutating Asn2 or Asn6 to Leu2 or Leu6 was shown to reduce the rate of aggregation. A mutation of Asn3 to Leu3, however, seemed to drastically accelerate aggregation. This research lays an important groundwork for future studies into the driving forces of peptide aggregation.

# Introduction:

*Protein Aggregation*

Peptides are among the most fundamental building blocks of life; their structural and enzymatic functions are what make even the simplest organisms possible. The amino acid sequence, known as the primary structure, and the initial folding of the chain into helices or pleated sheets, known as the secondary structure, are both vital to the function of the polypeptide. The tertiary and quaternary structures, or the way that peptides fold and assemble into larger proteins, are even more important to protein function. This is because the three-dimensional shape is a major determinant of how the protein works. Misfolding can lead to a broken or ineffective version of the molecule that is unable to perform its normal tasks, causing disease or death of the organism. Similarly, the aggregation and deposition of polypeptides can be highly dangerous, and has been linked to diseases including Alzheimer's, Huntington's, type II diabetes, and transmissible spongiform encephalopathies such as Mad Cow disease.

There are many different forms of protein aggregation. Technically, any protein can be made to aggregate through certain processes. For example, increasing the salt concentration can force the native folded proteins to amass together and precipitate out of solution. Such precipitates are easily dissolved, however, and cannot be truly classified as protein aggregates (1). True protein aggregation involves the cohesion of partially folded polypeptides into a large insoluble mass. Depending on the proteins involved, this mass can either be random and disordered or highly organized (1).

One common protein aggregate is known as an "inclusion body". Inclusion bodies are disordered aggregated structures that are generally produced by the over-

expression of proteins in cells. Although they are most commonly observed when a foreign protein is expressed in a bacterium, inclusion bodies can be found in many different cell types, from prokaryotes to eukaryotes (2). The general mechanism of aggregation is believed to involve intermolecular interactions between partially folded proteins. Essentially, as the peptide chain folds, it reaches a point where its aggregation potential is the highest. If there is a high concentration of similarly folded peptides, an interaction is more likely. Once the proteins begin to assemble together, they create an insoluble aggregate (1). This aggregate is considered to be disordered because it lacks the highly ordered superstructure of amyloid fibers. However, there is still a significant structural factor in the assembly of inclusion bodies; like amyloids, they have been found to form stacks of β-sheets. Furthermore, it has been shown that inclusion bodies can act as nucleation seeds. Once a small aggregate has formed, it will provide a location for other partially folded proteins to latch on and increase the size of the body. In essence, the presence of an inclusion body can induce further aggregation (2).

There are both beneficial and detrimental aspects of inclusion body formation. Such aggregates can be useful in the purification of peptides from transformed bacteria. Since the over-expressed proteins have gathered together into an insoluble mass, they can easily be separated from the rest of the cell. Furthermore, due to the overproduction of the single polypeptide and the ability of the inclusion body to induce further aggregation, the proteins found in the aggregates are generally homogenous and pure (1). Despite this potential benefit, there are still many negative effects of inclusion body formation. Because the proteins are in partially folded states, it can be difficult to return them to their native forms. Since misfolded proteins often lose their function, the researcher's

ability isolate functional proteins from the cells may be limited. Furthermore, if the peptides aggregate in an irreversible manner they cannot be properly studied (1, 2).

### *Amyloids*

Perhaps the most studied protein aggregates in recent years have been amyloid plaques. Amyloids are highly ordered proteinaceous masses that form into fibrils approximately 80 to 100 angstroms in width. They were first discovered by Rudolph Virchow in 1854. Believing the fibers to be carbohydrate based, he dubbed them "amyloids". It was later discovered that the amyloids were actually made of proteins (3). Since that time, extracellular amyloid plaques have been found to be associated with many human disorders, most notably Alzheimer's disease.

For a protein aggregate to be considered an amyloid, it must fulfill three requirements: the formation of fibers visible in electron microscopy; the ability to bind to the dye Congo Red; and high β-sheet content in the structure (1). While the formation of fibers represents the important superstructure associated with disease plaques, the β-sheet content of the secondary structure is also significant. In an amyloid, much of the normal secondary structure converts to β-sheets. These β-sheets stack together along the width of the fibril, forming what is known as the "cross-β structure" (3). The sheets themselves run parallel to the fibril axis, while the β-strands are perpendicular to it. Furthermore, some evidence now shows that the β-sheets form a helix, twisting around a common axis (4). These characteristics are common to all amyloids, regardless of the precursor proteins from which they were formed.

As the amyloid grows, its kinetics appear to be sigmoidal; there is a lag phase, followed by a rapid growth, eventually reaching a plateau. This is consistent with nucleation processes (5, 6). As with inclusion bodies, the amyloid aggregate forms more quickly when it has a preformed oligomer to attach to. To make the initial dimer, the peptides first begin to form β-sheets with each other. Once they are held together, the β-strands are open for further interactions. More peptides can now join the growing β-sheet, forming an oligomeric structure. These oligomers then interact with each other, creating the fibrils characteristic of amyloids (6). The initial β-sheet may be slow to form but, once it has, it becomes easier for other peptides to bind, thereby increasing the elongation rate. Furthermore, studies have shown that some amyloidogenic peptides have an "anti-chaperone" activity. A segment of the Alzheimer's β-peptide, known as Aβ(25-35) has been observed to actually break the native conformation of the unrelated protein firefly luciferase. Through its electrostatic interactions, the peptide was able to induce aggregation and even damage the proteins that remained in solution (7). This anti-chaperone characteristic may be important to amyloid formation. It is possible that some amyloids may be able to alter the native states *in vivo*, allowing for further aggregation.

Such amyloid formation becomes a problem because the aggregates formed are insoluble and unlikely to break apart. This is to the relative stability of the β-sheet structure compared to other secondary structures (8). While it is thermodynamically favorable for random coils or helices to coalesce together, these interactions are not as favorable as the formation of β-sheets. Therefore, once two proteins have aggregated into a β-sheet dimer, it requires energy to separate them into any other form. Since there

is no natural reason to break apart, the peptides remain aggregated, and eventually build up into an amyloid fibril. In fact, this thermodynamic preference can be even stronger than the favorability of the native globular state of the protein (8). This is why cells often use chaperones to ensure the proper folding of the peptide chains; without these molecules to guide the structural formation, aggregation would be a more common event (1, 9).

A major classification of amyloidogenic proteins is the "proteinaceous infectious particles". Better known as prions, these are proteins whose natural conformation has been altered, often allowing them to aggregate into amyloid-like fibers. The aggregates are insoluble, and are resistant to proteases (1). Several diseases known as "transmissible spongiform encephalopathies", including Mad Cow disease and Creutzfeldt-Jacob disease, have been attributed to prions (1, 10). What makes prions notable from other aggregating peptides is that they are known to catalyze the alteration of other molecules' native conformations. Essentially, if a protein becomes a prion, it has the ability to alter the shapes of similar proteins, turning them into prions. This function is usually very specific for self-aggregation; a prion of a certain protein will only affect molecules of that protein (11). Because of this, prions are known as an epigenetic factor, meaning they can be inherited without altering the genome (9). This is also why such spongiform encephalopathies are transmissible. If a healthy organism is infected with the prion protein, its own proteins may be altered, causing aggregation and the formation of the typical plaques.

Interestingly, prions may not always be detrimental. Because evolution seems to have conserved prion-forming domains, it would seem that they have some benefit (11).

It is possible that cells use this function to deal with over-expression or environmental stresses. Also, the fact that each prion will only affect very specific proteins keeps the function in check, and allows for several of such systems to be in place in a single cell.

### *Cytotoxicity*

There are several reasons why amyloids are dangerous in cells. Clearly the insoluble deposits can have a reducing effect on the total concentration of protein in solution. This could potentially strip the cell of vital enzymes or structural peptides. More important, however, seems to be the effects that amyloids have on cell membranes. These surfaces form an interface along which β-sheets can assemble, creating amyloid fibers near the edge of the cell (12). Interestingly, it is not the mature fibrils, but rather the oligomeric structures that are believed to be cytotoxic to the cells (6, 13). The mechanism for such cytotoxicity involves the formation of pores in the cell membranes, making the membrane permeable to some ionic molecules. This is especially important in neurological diseases such as Alzheimer's, where uncontrolled pores could destroy the ion potentials needed for nerve cell function (13, 14).

The formation of such non-native transmembrane channels can be observed in bacterial pore-forming toxins, which are proteins used to disrupt the ion balance of target cells. These pores are created using a series of β-strands in a ring known as a β-barrel (15). Unlike normal cellular β-barrel channels, they are assembled from several separate peptide chains instead of a single continuous peptide (14). Such structures can penetrate the cell membrane, connecting the cytoplasm and the extracellular environment and allowing ions to flow freely in and out of the cell (15).

Similarly, β-sheet aggregates have been shown to puncture cell membranes. Like in pore-forming toxins, it has been demonstrated that separate peptides can aggregate together to successfully create a channel out of β-strands (14). Moreover, these β-barrel structures may assemble together to form even larger pores. This ability has been observed particularly in peptides with repeats of hydrophobic amino acids, glycine, and serine. It was also determined that the dye Congo Red could inhibit this formation. This is strong evidence that the pore is being created by peptides aggregated into oligomers, since the dye is known to bind to amyloids (14). Similar experiments have shown that aggregate-formed pores are capable of allowing calcium ions and reactive oxygen species into the cytoplasm, which could be highly detrimental to a living cell (13).

In order to form a dangerous pore in the cell membrane, the peptides must first penetrate the membrane. This process is believed to be related to the natural assembly of the β-sheet structures along the edge of the cell due to van der Waals interactions between the peptides and the membrane lipids (12, 16). Peptides in a random coil formation begin to form very disordered aggregates in solution. Upon contacting the phospholipids of the cell membrane, the aggregate reassembles into β-sheets. With hydrophobic regions now exposed along the sheet, it naturally begins to insert itself into the membrane in order to access the hydrophobic tails of the phospholipids. Once inside the membrane, the β-sheets continue to interact, assembling the β-barrel structures and creating the membrane channels (14). A similar process has been shown to occur with the Alzheimer's amyloid β-peptide, known as Aβ. Once cleaved from the Alzheimer's precursor protein, Aβ finds and associates with "lipid rafts", or clusters of specific membrane lipids. Upon interaction with the lipid raft, Aβ changes its conformation from

α-helix to β-sheet, and forms the starting point for aggregation (17). From there, the peptide can penetrate the cell membrane, disrupting vital neural ion potentials.

Fortunately, in at least some cases the cells can be saved even after the oligomeric species have formed ion channels, as long as the damage has not been too severe (13). This allows for the possibility of treatment for amyloidogenic diseases. Since more and more diseases are being attributed to protein deposition, including Alzheimer's, Huntington's, spongiform encephalopathies, Parkinson's, and Type II diabetes, treatments for peptide aggregation are being heavily researched (15, 18).

### *Modulating Factors for Aggregation*

The effects of many small molecules, ranging from ions to other peptides, on the aggregation of Alzheimer's Aβ have been studied. Some, particularly glycerol and metal cations, were found to actually accelerate the aggregation process (19). On the other hand, monoclonal antibodies that target the Aβ peptide have had interesting effects on Alzheimer's disease. While they have not been shown to affect the amyloid plaque, there is evidence that antibodies can reduce the memory impairment associated with the disease (20). This is possibly due to the binding of the antibody to the soluble peptides. Since it is believed to be the oligomeric species that cause cytotoxicity, and not the mature fibrils, the antibodies may inhibit the effects of the oligomers without actually stopping the formation of the aggregated plaques.

Interestingly, modified peptides have proven to be a rather effective means of inhibiting aggregation. By introducing small peptides that aggregate preferentially with the target protein, amyloid formation can be prevented. Fragments of Aβ that can inhibit

the native aggregation have been identified, and have been modified to increase their usefulness. In particular, a cholyl-LVFFA-OH made with D-amino acids, instead of the natural L-amino acids used is organisms, is capable of effectively blocking the amyloid formation while remaining stable in a biological environment (21). This an important factor in designing an inhibitory peptide: not only must it be effective in reducing aggregation, but it also must be stable in cells so that it has time to function.

Another effective Aβ inhibitor has been designed to block the KLVFF region of the native peptide; it was created using alternating natural amino acids and α,α-disubstituted amino acids (22). These modified amino acids have the side-chains and the hydrogen replaced with various larger groups. Like the native peptide, the ααAA peptide is able to form β-sheets with other Aβ molecules. The extra side groups, however, block the β-sheet formation on one edge. Since other peptides cannot interact with this modified edge, the ααAA peptide effectively caps the sheet, and the aggregate loses its ability to grow. This design also ensures that the modified peptides will not self-aggregate, since they cannot interact with each other (22). Overall, such modified peptides have been shown to be effective ways to specifically target Aβ and to inhibit the formation of amyloids.

Since protein deposition may be detrimental to cells, it is not surprising that evolution has selected against peptides with high aggregation potentials (23). Generally, hydrophobic residues are kept within the protein's core. Since the hydrophobic regions will tend to come together to avoid the aqueous environment, this encourages proper folding, while keeping such regions away from other proteins. Similarly, β-sheets are usually protected internally, because external β-strands have the potential to interact and

form aggregates. Moreover, the β-sheets are often twisted into a β-barrel structure. Having no open ends, the sheets cannot connect to other peptides. In the cases where there is an external β-sheet, there is usually a loop or helix to protect it from extraneous interactions. In many polypeptides, the primary structure of the peptide has been designed to avoid β-strands. Such segments will contain glycines or prolines, which discourage β-sheet formation. Peptides rarely have alternating polar and non-polar residues, which tend to form amyloidogenic regions (23). Similarly, adjacent regions of proteins tend to have a varied amino acid composition, which makes aggregation less likely. Interestingly, some mutations in wild-type peptides have been predicted to further reduce aggregation potential, indicating that evolution has not completely protected the proteins (24). Perhaps fully protecting the peptides from aggregation would interrupt certain vital functions, such as the ability to interact with other peptides via glutamine and asparagine rich regions (11, 25). Alternatively, it is possible that the proteins have simply been protected enough to prevent aggregation within a reasonable life span (24).

### *Aggregating Sequences*

One of the major reasons that proteins needed to evolve specific defenses against amyloidogenic species is that aggregation seems to be a generic property of peptide chains. There is no specific sequence of amino acids that causes protein deposition. Rather, many different and unrelated sequences are capable of forming aggregates (23). Originally this quality was believed to be solely found in hydrophobic regions, but it has now been observed in a variety of hydrophilic amino acid sequences as well. Even the size of the amyloidogenic peptides is inconsistent; for example the commonly studied

forty-two amino acid Aβ42, compared to the fibril forming tetrapeptide DFNK (18). However, the most important factor for aggregation propensity seems to be the ability of the peptide to form β-sheets. Certain amino acids favor the β-strand structure more than others, based on the steric interactions between the amino acid side chain and the backbone (26). However, a higher likeliness to form a β-sheet does not necessarily indicate a higher rate of aggregation. Instead, peptides that can aggregate slowly but steadily build more fibrils. It is important for the peptide to be able to convert to β-sheets, but these must polymerize at a rate that will allow the super-structure to form (25). Just as mutations affecting the ability to form β-sheets have a significant effect on the ability of the peptide to aggregate, so do those that affect hydrophobicity and charge. These three factors can be used to estimate the aggregation propensity of a given polypeptide chain (24).

Among the various amyloidogenic peptide sequences, hydrophobic regions remain amongst the most important and best studied. This is because the amyloidogenic properties of Alzheimer's Aβ peptide are based on its highly hydrophobic nature. There are several hydrophobic segments in Aβ. Among these, the central region has often been considered to be the most important, but others including the C-terminal region have also been shown to be significant. Mutating residues in these regions to hydrophilic amino acids causes a reduction in aggregation rates, although certain mutations have greater effects than others (27). Such mutations will also affect the thermodynamics of the aggregation, reducing the overall amount of amyloid formed (28). In fact, the hydrophobicity of Aβ is so important that randomly mutating the hydrophobic regions of the peptide with different hydrophobic amino acids does not disrupt the ability to

aggregate. While this will cause some minor differences in kinetics and fibril structure, any of the hydrophobic residues can be replaced with any other, and the peptide will still maintain its amyloidogenic properties. Essentially, it is the hydrophobicity, and not the specific amino acid sequence, which gives Aβ its ability to form aggregates (29). Therefore, if the hydrophobicity is increased by a mutation, then the aggregation rate and the thermodynamic favorability of the reaction will also increase (28). There are still other important factors, however, such as β-sheet propensity and overall peptide charge. Altering these features affects the amyloid formation, and is responsible for the differences in kinetics and fibril morphology seen when random hydrophobic mutations are made (30). Overall, the hydrophobicity of the amino acid sequence is what causes Aβ to aggregate, although β-sheet propensity attenuates the rate.

Many hydrophobic peptides other than Aβ have demonstrated tendencies to aggregate. The hexapeptide NFGAIL, from the human islet amyloid polypeptide (IAPP), is highly hydrophobic, and has been shown to form amyloid fibrils (31). Hydrophobic forces will cause this peptide to form a disordered association with an oligomeric aggregate. Such an interaction is highly favorable because it reduces the amount of water molecules around the hydrophobic residues. After the random coil NFGAIL has associated with the ordered oligomer, it will shift its structure to form a β-sheet with the rest of the aggregate, further reducing the overall energy state (31).

Other hydrophobic aspects of peptide chains can aid in aggregation, even in predominantly hydrophilic peptides. Regions of alternating hydrophobic and hydrophilic residues favor β-sheet formation, and therefore are often indicative of amyloidogenic peptides (29). Similarly, aromatic residues are important for amyloid formation.

Aromatic ring structures have been shown to stack together in aggregates of the hydrophilic peptides DFNKF and GNNQQNY. This provides a stabilizing force for the aggregate, allowing the amyloid to grow (10, 18, 32).

While generic hydrophobic regions are sufficient to cause aggregation, there are some specific amino acid sequences that have been implicated as well. Using mutants of the peptide STVIIE, certain basic patterns of amyloidogenic amino acid sequences have been found. These patterns have been shown to correspond to actual aggregating peptides. While it is unlikely that the patterns can be used to describe all aggregating peptide regions, it is an important step in identifying polypeptides with the potential to form amyloids (25).

An important type of hydrophilic region that has been found to cause peptides to aggregate involves high glutamine and asparagine content. Regions rich in Q or N have been observed in several aggregating proteins, including the yeast prion Sup35 and the Huntington Protein. In fact, Q/N-rich domains can be found in a large variety of proteins in organisms ranging from bacteria to eukaryotes. Such domains appear in peptides far more commonly than can be explained by simple random chance, implying that these aggregation-prone regions are actually selected for by evolution. The Q/N-rich regions probably perform an important cellular function, likely involving protein-protein interactions or self-aggregative properties. Interestingly, thermophilic organisms tend to lack these domains, possibly due to the increased rate of aggregation at higher temperatures. Prions form when a polypeptide containing such a region is not properly folded (11). Normally, cells use chaperones to ensure that such domains remain

innocuous. For example, chaperone molecules have a significant effect on whether the yeast prion Sup35 folds properly or becomes an amyloidogenic prion (9).

### *Sup35 Fragment GNNQQNY*

A fragment from the Sup35 protein, the heptapeptide GNNQQNY (Figure 1A), has been identified as a region that can aggregate into amyloid fibrils. Many computer simulations have been performed to study the aggregation process. In solution, a single molecule of GNNQQNY is predicted to take a random coil structure. Although the peptide is known to have a high aggregation propensity, the formation of an α-helix is actually more thermodynamically stable than a β-strand. This energy barrier for β-strand formation slows aggregation, which explains why a nucleation seed helps speed the process (33). When two GNNQQNY strands interact, they will transiently form a β-sheet. The sheet tends to take on an anti-parallel structure for the formation of the initial dimer, simply because the anti-parallel sheet requires less energy than the parallel sheet (10). Because a β-sheet has now formed, however, the energy barrier for the next strand is lessened. When the third peptide attaches, a parallel alignment is now favored (5). This effect is due to the C-terminal tyrosines; the aromatic rings stack upon one another, stabilizing the parallel structure (5, 32). The formation of an anti-parallel sheet is allowed, but is less stable and therefore more likely to break apart. Thus, the initial dimerization of GNNQQNY is based on an anti-parallel β-sheet, while further polymerization occurs via parallel alignments (5, 10).

As the parallel β-sheet forms, they begin to take on the "dry, steric zipper" that is characteristic of aggregates formed by GNNQQNY and NNQQNY. This structure is

unusual because it involves a dry interface free of water molecules, but is constructed by hydrophilic Q and N residues (34). The steric zipper is possible due to the amino acid structure of the peptide. According to computer models, there are many interpeptide and intrapeptide hydrogen bonds formed in addition to the backbone hydrogen bonds of the β-sheet. Based on the general structure of the β-sheet, the side-chains at positions 1, 3, 5, and 7 point away from the dry interface, while positions 2, 4, and 6 point into it. The Asn6 forms an intrastrand hydrogen bond with Gln4, while Asn3 hydrogen bonds with Gln5. Meanwhile, the asparagines at positions 2, 3, and 6 form interstrand hydrogen bonds with their counterparts in the peptides above or below them. These bonds effectively reduce the amount of water that can interact with the hydrophilic residues, which is entropically favored, since it releases water molecules from the interaction (33). Without the water, the β-sheet can behave as if it has a hydrophobic face, and will interact with another sheet to form a dry interface. The steric interactions between the two sheets further favor this interaction, creating the "dry, steric interface". Essentially, the Asn2, Gln4, and Asn6 of two peptides associate together in opposite directions. This interaction keeps the two β-sheets together, and aids in the aggregation of the peptides (34).

### *Goals of this Study*

While many computer simulations of GNNQQNY or NNQQNY aggregation have been performed, few true chemical mutagenesis reactions have been reported. Computer simulations are very useful, but it is still important to verify the models using real chemistry. To that end, the peptides GNNQQNY and NNQQNY, the leucine mutants

GLNQQNY, GNLQQNY, GNNQQLY, LNQQNY, NLQQNY, and NNQQLY, and the valine mutants NVQQNY and NNQQVY were synthesized (Figures 1 and 2). Experiments were performed on these peptides to elucidate the importance of the specific amino acids and to determine the relative importance of steric factors and electrostatic interactions such as hydrogen bonding. The aggregation of various peptides was studied, and an important groundwork was laid for future research into the driving forces of peptide aggregation.

## Materials and Methods:

*Solid Phase Peptide Synthesis*

Fmoc solid phase peptide synthesis was performed manually to make many of the peptides. 0.2mmol of Fmoc-Tyr-(tBu)-Wang Resin was used for each peptide. The amino acids were added in a five-time excess, using 1mmol each. The resin was first washed and swollen with dimethylformamide (DMF) for twenty minutes. It was then deprotected using 20% Piperidine in DMF twice, each for ten minutes. The amino acid to be added was activated using 0.9mmol 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 1.4mmol of diisopropylethylamine (DIEA) and 2mL of DMF. After two minutes, the reaction vessel was drained of DMF, and the activated amino acid was added and allowed to react for twenty minutes. The reaction vessel was drained, and the resin was washed with DMF to remove excess reactants. This process was repeated until all amino acids were added. A final deprotection was performed to remove the Fmoc protecting group from the last amino acid, and the resin was allowed to dry.

Automated solid phase peptide synthesis was performed on the Protein Technologies Inc. Tribute Peptide Synthesizer, using similar procedures. The reactions were run on either 0.1mmol of resin or 0.3mmol of resin scales, still using five equivalents of amino acid. The amino acids were either prepackaged in vials with HBTU or were assembled in the lab. DMF could be replaced with N-methyl-2-pyrrolidinone (NMP), and DIEA could be replaced with 0.40M N-methlymorpholine.

*TFA Peptide Cleavage*

Peptides were cleaved off the solid support resin using a trifluoroacetic acid (TFA) solution. Generally for the peptide GNNQQNY and its leucine and valine mutants, water and triisopropyl silane (TIS) were sufficient scavengers to aid in the removal of protecting groups. The solution was made to be 95% TFA, 2.5% water, and 2.5% TIS, using about 1mL for every 100mg of resin and peptide. For peptides with a larger array of protecting groups, Reagent K was used. This solution contained 85% TFA, 4% water, 4% thianosole, 4% phenol, 2% 1,2-ethanedithiol (EDT), and 1% TIS. In both cases the reaction was allowed to run for one hour before draining the TFA solution, now containing the peptide, into a beaker.

Normally, an ether precipitation would be performed. This involves adding cooled diethyl ether into the solution to cause the peptide to crash out. The solution would then be centrifuged, and the ether layer would be poured off. This process would be repeated, and the solid would be allowed to dry.

Since GNNQQNY and its mutants have no charged residues, the peptides failed to crash out of solution in the presence of ether. Instead, a cooled ten-to-one hexane-to-ether mixture was added to the TFA solution and mixed. The hexane layer, now containing many of the hydrophobic protecting groups, was then be poured off, while the TFA layer was evaporated down, and then diluted with HPLC buffer A.

*HPLC Purification*

The peptides were purified via High Pressure Liquid Chromatography (HPLC), using a Waters PrepLC System. The HPLC used two buffers: buffer A was 95% water, 5% acetonitrile, and 0.1% TFA, while buffer B was 95% acetonitrile, 5% water, and 0.1% TFA. Outflow was monitored using an ultraviolet and visible spectrophotometer, measuring absorbance at 220nm, for the peptide bond absorbance, and 280nm, for the tyrosine ring. The peptides were run through a 250x30.00mm 10 micron column over a fifteen minute gradient of increased buffer B. GNNQQNY and its mutants usually came off the column in sharp, distinct peaks when a 100%A, 0%B to 40%A, 60%B gradient was used. The peaks were collected, and samples of the fractions were sent for mass spectrometry to verify the presence of the peptide. The rest of the fractions were combined and evaporated down. The peptide in water was then frozen and lypholized.

*LCMS*

Liquid Chromatography/ Mass Spectrometry (LCMS) were performed using the Micromass LCT system. This technique was used to assess the purity of the peptides after HPLC.

*NMR*

Proton Nuclear Magnetic Resonance Spectrometry (NMR) was performed using the Varian Gemini 400MHz Nuclear Magnetic Resonance Spectrometer. This was used to assess the identity and purity of peptides and thioflavin T.

*Peptide Solution Preparation*

The peptide solutions were prepared by dissolving solid peptide in dimethyl sulfoxide (DMSO), usually keeping the final DMSO content at 0.7%. The solution was then diluted with buffer and then sonicated for twenty minutes. To remove any aggregates, which could act as potential nucleation seeds, the solution was filtered through Millipore Ultracel YM-3 3kDa centrifuge filters. The concentration was determined by measurement of absorbance at 280nm, and the solution was diluted to the desired concentration with buffer, keeping the DMSO content constant.

*Turbidity Measurement*

Turbidity measurements were taken on the PerkinElmer Lambda25 UV/VIS spectrometer to monitor the aggregation kinetics of the peptides. Absorbance of peptide solutions was monitored at 400nm for several days at 25°C, either with or without stirring.

*Monomer Assay*

Post-aggregation monomer concentrations were measured. The aggregate solution was centrifuged through 3kDa filters to remove the aggregated peptides. The concentration of peptide remaining in solution was then measured via absorbance at 280nm.

*Fluorescence Spectrometry*

Peptide aggregation was measured using Spex Fluorolog Fluorescence Spectrometer. A peptide solution was prepared with 2mM thioflavin T. The solution was excited at 440nm, and fluorescence was measured at 485nm.

*Recrystallization of Thioflavin T*

Thioflavin T (ThT) was recrystallized to increase its purity. The ThT was dissolved in the minimal amount of warm water. The solution was cooled and filtered. The purified ThT crystals were collected and dried.

# Results:

*Synthesized Peptides*

The peptides NNQQNY, LNQQNY, NLQQNY, NNQQLY, NVQQNY, NNQQVY, GNNQQNY, GNNQQLY, GLNQQNY, and GNLQQNY were synthesized using Fmoc solid phase peptide synthesis (Figures 1 and 2). Each was then purified on the HPLC. The purity of the peptides was assessed using proton NMR and LCMS. NMR data for GNNQQNY in $D_2O$ (Figure 3) indicates a high purity of the peptide, as there is no discernable impurity in the spectra. Similarly, the LCMS data for the peptides demonstrate high purity. A comparison the areas of the major peptide peaks to the impurities show GNNQQNY to be approximately 80% pure, and GLNQQNY to be 90% pure (Figure 4). GNLQQNY, GNNQQLY, and NNQQNY and its mutants each have a purity of 95% or greater.

*Turbidity Measurements*

Kinetic studies of the peptide aggregations were performed by monitoring the absorbance of the solution at 400nm. The first peptide studied was NNQQNY at 0.30 mg/mL. In a 10mM phosphate buffer at pH 7.4 and 25°C, very little change in absorbance was observed (Figure 5A). The kinetic study did show a general rise in absorbance, which could potentially indicate aggregation, but the overall change was very small. To test if the peptide was responsible for the slight increase, the buffer was run on the spectrometer alone (Figure 5A). No increase in absorbance was observed, indicating that the peptide was the factor that caused the increase in absorbance.

Because the crystal structure of NNQQNY had been determined using high zinc cation concentrations (34), the effect of zinc on the aggregation kinetics was tested (Figure 5B). A HEPES buffer with 10mM $ZnCl_2$ at pH 7.5 was prepared, and the peptide was added. After reacting for over fifteen hours, there were no signs of aggregation.

In an effort to speed up the kinetics of the aggregation, high salt concentrations were added to the solutions. NNQQNY in 10mM phosphate buffer with 300mM NaCl at pH 7.4 and 25ºC still showed only a slight increase in absorbance (Figure 5C). The temperature was raised to 30ºC in a further attempt to increase the rate (Figure 5D). Unfortunately, the rise in absorbance was still small until near the 30 hour mark, at which time an abrupt increase was observed.

Because the reaction time was so long and the increase in turbidity was so low for NNQQNY, the heptapeptide GNNQQNY was used instead. In general, GNNQQNY showed a significantly higher change in turbidity compared to the hexapeptide counterpart. Several turbidity measurements of 0.30 mg/mL of GNNQQNY in 10mM phosphate buffer at pH 7.4 and 25°C were performed. Without stirring, the graph appeared jagged due to a lack of homogeneity in solution. This graph could be corrected to show the general rise in absorbance (Figure 6A). When stirred, the kinetics of the aggregation could be seen much more clearly (Figure 6B), but continued to be jagged due to peptides crashing out of solution. With attenuated stirring, a much smoother kinetic graph was observed (Figure 6C). Still, the reaction took nearly 100 hours to finish, which is impractical for comparing several peptides.

*Determination of Ideal Aggregation Conditions*

This long time course indicated that the conditions for the peptide aggregation were not ideal. In order to optimize the reaction, a series of aggregations was performed with varied pH and salt concentration, measuring the post-aggregation monomer concentrations. The buffers used were: 0.01M HCl (at pH 2), 10mM sodium acetate, 10mM boric acid, 10mM Tris(Hydroxymethyl) Aminomethane (at pH 4, 6, 8, and 10), 10mM NaPi (at pH 8), and 10mM KPi, 300mM NaCl (at pH 7.4). The first series of reactions is summarized in Table 1. It indicated that the Tris buffer at pH 8 and the high salt concentration both increased the reaction rate.

This was not entirely conclusive, however, because each sample had different starting concentrations of the peptide and DMSO. Therefore, a second series of reactions was performed, as summarized in Table 2. In this series, the initial concentration of each sample was set to 1mM, or 0.67 mg/mL, and the DMSO concentration was kept constant at 0.7%. The samples were first run for two days without agitation, and then for another two days with agitation. Unfortunately, the samples were not agitated equally, which makes the data unreliable. In pH 4 buffer over half of the original peptide had aggregated, but this is possibly due to extra agitation time. The only corresponding data was in the high salt buffer, which still showed one of the higher aggregation percentages.

Although the pH series proved mostly inconclusive, it was determined that increasing the salt concentration should increase the aggregation rate. The next step was to isolate the ideal concentration for kinetic studies. The turbidity kinetics of several high concentrations of GNNQQNY in 10mM KPi and 300mM NaCl at pH 7.4 were monitored. Amongst 0.63 mg/mL, 0.90 mg/mL, and 1.5 mg/mL, the lowest

concentration was found to have the clearest trace (Figure 7A). The two higher concentrations seemed to aggregate too quickly, because of which the lag phase in the sigmoidal curve was not apparent. Post-aggregation monomer measurements of the samples showed that 56% of the 0.63 mg/mL GNNQQNY, 58% of the 0.90 mg/mL, and 71% of the 1.5 mg/mL had aggregated (Table 3). The ideal concentration was further isolated by observing the kinetics of 0.49 mg/mL, 0.55 mg/mL, and 0.60 mg/mL (Figure 7B). The 0.60 mg/mL sample had the clearest sigmoidal kinetics and was determined to be the best concentration for further studies.

### *GNNQQNY Mutant Turbidity Studies*

After the optimal conditions for aggregation were determined, turbidity measurements of the mutants of GNNQQNY were performed. The first series measured the aggregation of GNNQQNY at 0.60 mg/mL, GNNQQLY at 0.60 mg/mL, and GNNQQLY at 0.50 mg/mL (Figure 8A). An effort was made to test GNLQQNY, but the initial concentration of the peptide was too low. Although the mutant was properly dissolved in solution, it was nearly entirely filtered out by the 3kDa filters. The 0.50 mg/mL GNNQQLY solution did not show significant aggregation, while the 0.60 mg/mL GNNQQLY showed a drastically reduced aggregation rate. The post-aggregation monomer concentrations of this experiment showed that about 60% of the wild type peptide aggregated. About 57% of the GNNQQLY at 0.60mg/mL aggregated, while only 18% of the peptide at 0.50mg/mL aggregated (Table 3). It is important to note, however, that these monomer measurements were taken several weeks after the kinetics

experiments. Therefore, it is difficult to tell when the mutants reached thermodynamic equilibrium, or if they had even reached it by the time the measurements were taken.

Another series of mutants was tested, with GNNQQNY, GLNQQNY, and GNNQQLY at concentrations of 0.60 mg/mL (Figure 8B). Once again GNLQQNY was not able to be used due to being filtered out of solution. GNNQQLY did not show significant aggregation on its kinetic trace, and only 33% of the monomers aggregated. The trace of GLNQQNY showed some aggregation, but at a significantly reduced rate. Its monomer concentration, taken several weeks after the kinetic experiment, showed that 77% of the peptide aggregated (Table 3). Again, it is difficult to determine at what point the reaction reached equilibrium.

To determine the validity of the results, it was important to measure their repeatability. This was done by performing a triplicate reaction of GNNQQNY at 0.50 mg/mL (Figure 9). Each sample was treated identically, with the same concentrations in the same 10mM KPi, 300mM NaCl buffer. The results of this experiment showed two kinetic traces that were similar, with the $t_{1/2}$ differing only by approximately two hours. The third sample did not show any sigmoidal curve or a significant change in absorbance. However, the post-aggregation monomer concentrations for this series of experiments shows that the peptides aggregated equally, each having about 50% of the monomers aggregate (Table 3).

### *Fluorescence Data*

Along with turbidity studies, fluorescence spectrometry was used in an effort to monitor peptide aggregation. A proton NMR of the recrystallized thioflavin T in $D_2O$

indicated a pure product (Figure 10). The GNNQQNY aggregation experiment run on the spectrometer showed no evidence of aggregation. To determine whether ThT binds to GNNQQNY aggregates, three different concentrations of aggregated samples were tested. Both the 0.63 mg/mL and 0.90 mg/mL samples showed significant increases in intensity over the ThT alone in buffer (Figure 11). The fluorescence of the 0.63 mg/mL sample was over twice as intense, and the 0.90 mg/mL sample was three times as intense. The 1.50 mg/mL sample of aggregated GNNQQNY, however, showed no significant increase in intensity. This could indicate that the highly concentrated sample aggregated differently, either achieving more mature fibrils or perhaps not forming true amyloids due to the high reaction rate.

## Discussion:

In order to test the importance of the specific amino acids in the amyloidogenic peptide GNNQQNY, several different mutants were synthesized. It was hypothesized that mutating the various asparagines would significantly alter the kinetics of aggregation, since these residues have been demonstrated by computer model (33) and crystal structure (34) to be involved in interstrand and intrastrand hydrogen bonding. Leucine is similar in size to asparagine (Figure 12A and B), although it has an SP3 bond character instead of the SP2 found in asparagine. The major difference between the two residues is that leucine is hydrophobic and will not form hydrogen bonds. Therefore, mutating the asparagines into leucines may further clarify the aggregation process. If the leucine mutants show an increased aggregation rate, then it is possible that the hydrogen bonds are not vital to the assembly of oligomeric structures. On the other hand, if the mutants aggregate more slowly, it may indicate that the interstrand and intrastrand hydrogen bonds are a major component of the aggregation process.

### *Optimization of Aggregation Conditions*

An important factor in the experimental process was determining the ideal conditions for the aggregation of the peptides. The first notable condition was agitation, which was shown to accelerate the aggregation process. This can be clearly seen from the pH optimization experiments (Table 2). The pH 4 buffer solution had no appreciable aggregation after two nights sitting motionless, but over 50% of the peptide aggregated after two nights of being shaken. Agitation was also important for turbidity

measurements. Stirring the samples on the UV/VIS ensured that the solution was homogenous and that an accurate assessment of the turbidity was being made.

Another factor that may have been important was the DMSO concentration. DMSO is known to keep amyloidogenic peptides soluble, which is why it was used to initially dissolve the peptides. Generally the concentration is kept low by using minimal DMSO and diluting the solution with buffer. Early in the tests, however, the DMSO percentage was not accounted for when diluting the peptide solutions to the desired concentration. Although the specific effects of different DMSO percentages were not tested, there was some concern that it may alter the aggregation kinetics. Therefore, in order to ensure that it would not be responsible for any differences in kinetics, the DMSO was kept constant at 0.7% for all samples.

Two series of experiments were performed to find the optimal pH for the peptide aggregation. Unfortunately, the pH tests were rather inconclusive. During the first series (Table 1), it seemed that pH 8 was ideal, since it showed the greatest peptide aggregation. This result was uncertain, however, because each sample had a slightly different initial concentration of peptide and DMSO. While this difference was small, it may have been great enough to affect the results. Therefore, a second series of tests was run, keeping the concentrations identical for each sample. This experiment turned out to be even less conclusive. This time, the peptides did not aggregate until they were agitated, but due to an error in the shaking process they were not agitated equally. Since agitation is so important to the peptide aggregation, any variability in the shaking would alter the results. Although neither experiment was completely conclusive, the tests did seem to indicate that the pH was not as important for the aggregation as other factors, such as

stirring. Therefore, it was determined that a pH of 7.4 would be acceptable for aggregation experiments.

Perhaps the most important condition turned out to be the salt concentration. Based on the turbidity measurements, it seemed clear that the peptides did in fact aggregate, but that the rate was too slow for efficient experimentation. Increasing the salt concentration sped up the reaction by forcing the aggregated peptides out of solution. This effect can be seen in the pH tests (Tables 1 and 2), in which peptides in high salt buffer aggregated more readily than most of the others. Through all of these conditions, it was determined that a buffer consisting of 10mM KPi and 300mM NaCl at pH 7.4 was optimal for the aggregation kinetics.

### *Aggregation Kinetics of NNQQNY and GNNQQNY*

Initially, the hexapeptide NNQQNY was intended to be the main control peptide. This did not work, however, because NNQQNY did not demonstrate a clear sigmoidal curve in the turbidity reaction (Figure 5). It did seem to aggregate, since the peptide solution increased in absorbance more than the buffer alone. This rise was very small, however, and could not be accurately studied. The heptapeptide counterpart GNNQQNY was then used instead. This peptide aggregated much more readily than NNQQNY (Figure 6), and it gave a much clearer kinetic trace. This result is very interesting; by simply adding a glycine to the peptide, the turbidity studies change drastically. Moreover, glycine residues are generally expected to break β-strands and reduce aggregation, but in this case it seems to be vital to aggregation.

It is strange that the addition of a glycine, an amino acid known to interfere with the formation of β-sheets, would aid in the peptide aggregation. Perhaps glycine is not the ideal residue to induce aggregation. Instead, an amino acid with a higher β-sheet propensity could potentially increase the rate even further. It would be very informative to create another mutant of GNNQQNY, replacing the glycine residue with a different amino acid. Such an experiment could elucidate the importance of the seventh residue.

A likely reason for the lack of turbidity with NNQQNY is that the peptides are simply not assembling into large enough structures. They may be forming β-sheets, but unless they are able to aggregate into the oligomeric superstructure the turbidity will not increase. It is possible that the glycine on the GNNQQNY provides an interaction point for other peptides, allowing the large aggregates to form. This can be seen in the crystal structure data, where the glycine residues seem to interact with other peptides, stabilizing the overall structure (34). A Congo red binding assay and circular dichroism would be useful in determining whether NNQQNY takes on a β-sheet structure in solution.

### *Leucine Mutants of GNNQQNY*

When the leucine mutants of GNNQQNY were tested, both GLNQQNY and GNNQQLY demonstrated lower aggregation rates than the wild type peptide (Figure 8). GNNQQLY also seemed to have less of its peptide monomers aggregate together than the wild type (Table 3). GLNQQNY, on the other hand, showed an increase in monomer aggregation over GNNQQNY. It is difficult to determine when the mutant peptides reached thermodynamic equilibrium, however, since the kinetic traces of these peptides did not show the final plateau and the monomer concentrations were measured several

weeks after the kinetic experiments. It seems that the Leu2 and Leu6 mutants both reduce the rates of aggregation, while Leu2 may still increase the total aggregate formed at thermodynamic equilibrium. This means that while GLNQQNY takes longer to aggregate, its equilibrium will have more aggregated material than the wild type. It is also possible that Leu6 will have a similar result if left to react for a long enough period of time.

Both of these peptides have mutations on the dry interface of the steric zipper. Changing Asn2 and Asn6 to Leu2 and Leu6 would be expected to interfere with the hydrogen bonding found in the wild type. It is reasonable, therefore, to expect that such mutations would reduce the aggregation rate. Even though the hydrophobic leucine residues would tend to prefer a dry environment, the loss of hydrogen bonding capacity may interfere with the peptides' ability to form the dry interface. While the lack of hydrogen bonding increases the total reaction time, the extra hydrophobicity may still cause more of the peptide monomers to aggregate at equilibrium. This would explain how the Leu2 mutant shows a decreased reaction rate, but an increased amount of aggregated material.

It is also possible that leucine's slight difference in shape, due to its SP3 character, inhibits the aggregation through steric interactions. Since it is not the exact same shape as asparagine, it may not fit properly into the highly ordered steric zipper. Unnatural amino acids could be designed that are more similarly shaped to asparagine, while still lacking the hydrogen bond capacity (Figure 12). The aggregation kinetics of a mutant peptide using such an amino acid could help to clarify the importance of the steric interactions. Still, this preliminary data seems to support the computer models and

indicates that the side-chain hydrogen bonds may be very important for the peptide aggregation. The experiments certainly need to be repeated, however, to ensure accuracy.

Unlike the Leu2 and Leu6 mutants, the third leucine mutant GNLQQNY seemed to drastically increase the aggregation rate. Unfortunately, the kinetics of the reaction and the monomer concentrations could not be monitored because the peptide crashed out of solution too quickly. This lack of hard evidence makes it difficult to draw a conclusion. The aggregate formed may not even be an actual amyloid-like structure. Instead, the peptide may have lost solubility in the high salt solution. Although this seems unlikely, since the very similar GLNQQNY was soluble, there is currently no evidence to study. Circular dichroism needs to be performed to verify the identity of the secondary structure and tell for certain whether or not the Leu3 mutant has formed a true aggregate.

Still, there is logic to the idea that the Leu3 mutant would increase the aggregation rate. Asn3 is located on the wet interface, and is generally involved hydrogen bonding interactions with water molecules. Leu3, being hydrophobic, would not favor a wet interface. Instead, it is possible that the leucines from different molecules will interact through hydrophobic forces. Such an interaction could potentially accelerate the aggregation process by destroying the wet interface and creating a new hydrophobic interface. This hypothesis could be tested by mutating Gln5 to a leucine or other hydrophobic residue, making GNNQLNY. Since Gln5 is also on the wet interface, mutating it to a hydrophobic residue may have a similar effect on the aggregation rate.

*Fluorescence Spectrometry*

In addition to the turbidity studies, fluorescence spectrometry was used to monitor the aggregation of GNNQQNY. This was done using the fluorescent dye thioflavin T, which was chosen due to its known ability to bind to amyloid fibrils formed by Alzheimer's Aβ. It became problematic, however, because it has not been proven to bind to all amyloidogenic peptides, including many of the smaller ones. Therefore, it is unclear how ThT it interacts with GNNQQNY. During the aggregation studies, no change in fluorescence was seen. This could mean that the dye did not properly bind to the aggregating peptide, or that the peptide did not aggregate at all. Since the solution was not stirred during the course of the experiment, it is possible that it simply failed to aggregate.

To test the ThT, known aggregated samples were examined for fluorescence in the presence of the dye. The lower concentrated samples fluoresced more than the ThT alone, but the highest concentrated sample did not (Figure 11). This seems to indicate that the dye binds to some form of the peptide aggregate, but for some reason does not bind to highly concentrated samples. It is possible that ThT will interact with intermediate oligomers of GNNQQNY but not the final fibril. Since the highly concentrated sample is likely to have reached this final stage more readily, it may not bind ThT. This hypothesis is not supported by the literature, however, which indicates that ThT only binds to mature fibrils. It is more likely that the highly concentrated sample aggregated so quickly that it could not form the regular superstructure of the

amyloid. Instead, the peptides may have simply massed together without forming fibrils for the ThT to bind to.

Another aggregation experiment needs to be performed, with stirring, so that the effect of ThT can be monitored over the course of the reaction. It would also be helpful to perform a positive control on the ThT by monitoring the aggregation of Aβ. It is known that the dye should bind to the Alzheimer's peptide, so this experiment would verify that the thioflavin T is functioning properly. Since the proton NMR indicates that the ThT is properly formed and pure, however, it is unlikely that there is a problem with the dye itself (Figure 10).

### *Reliability of Data*

Of course, it is always important to check the veracity of the data. Based on LCMS and proton NMR, the peptides were at least 80% pure, and were usually around 95% pure. This high purity makes it unlikely that the changes in turbidity or aggregation were due to the presence of contaminants. Also, the data proved to be fairly repeatable. During the triplicate GNNQQNY aggregation, two of the traces were similar, with half-times differing by only two hours (Figure 9). The third sample did not give a good kinetic trace. However, the peptide in this sample still seems to have aggregated, since there is a general increase in absorption early on. It is likely that there were problems with the stirring, which would cause the aggregates to precipitate out and would therefore alter the absorption. This is supported by the post-aggregation monomer concentrations, which show that all of the samples aggregated equally (Table 3). Such data indicates that the turbidity measurements do not always prove whether or not aggregates have formed.

Clearly, it is important to repeat experiments to verify the results. With enough repeated data, it should be possible to come to a reliable conclusion.

A major question that must be asked is exactly what sort of aggregate has formed. Turbidity measurements are not necessarily indicative of amyloid formation. Since studies of GNNQQNY aggregation have been published describing the amyloid formation, it is unlikely that any other sorts of aggregates are forming. Still, it is important to verify the identity of the aggregate using Congo red dye, circular dichroism, or microscopic studies of the fibrils.

### *Future Experiments*

There are still many experiments that can be performed to further elucidate the driving forces behind GNNQQNY aggregation. As previously stated, many of the experiments, especially with the leucine mutants, need to be repeated to ensure their accuracy. It would also be informative to make Gly1 mutants of GNNQQNY, particularly with amino acids of high β-sheet propensity, to study the function of this residue. NNQQNY should be studied in more depth, along with its leucine and valine mutants. Of particular interest would be monitoring the aggregation of NLQQNY, to see if it had a similar effect as the Leu3 mutant of GNNQQNY. More post-aggregation monomer concentration studies need to be performed. This technique is useful in studying the degree of aggregation for the peptide. Furthermore, it can be used to determine if NNQQNY is aggregating into β-sheets without forming larger oligomers. If the monomer concentration decreases in the reaction, then NNQQNY is in fact forming aggregates. It is also necessary to determine the identity of the aggregates using Congo

red or circular dichroism. This is particularly important in determining if GNLQQNY is accelerating amyloid formation or simply causing a different form of peptide deposition. If the Leu3 mutant did in fact create an amyloid so quickly, Asn3 would be a very interesting residue for further study.

In addition, new mutants should be made with unnatural amino acids. The amino acid in Figure 12C is a much better mimic of asparagine than leucine is. Due to its double bond, it has SP2 characteristics and is planar. Also, it has a similar size and shape to asparagine. Unlike the wild type amino acid, however, it cannot form hydrogen bonds. A GNNQQNY mutant with this amino acid could be used to examine the importance of the hydrogen bonding abilities while maintaining the steric characteristics. Similarly, the amino acids in Figures 12D and 12E can mimic the general size and shape of asparagine, but have different sized residues that could be give information of the steric allowance. It would also be interesting to create a peptide with a ring structure in the place of the Asn2 and Gln4 side chains, which would imitate the structure formed by hydrogen bonding interactions. Since it would be a fixed structure it may increase the aggregation rate.

## *Conclusion*

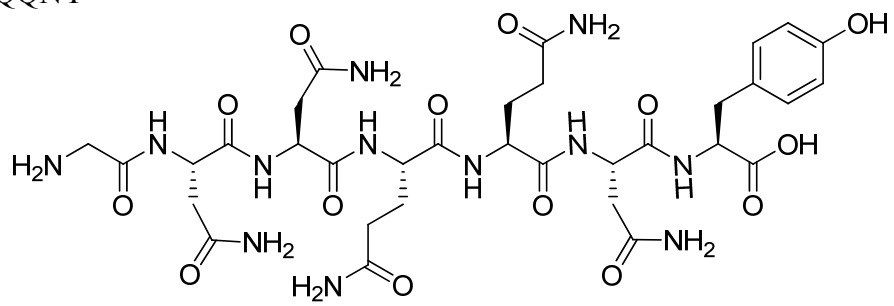
This series of experiments was an effort to study the driving forces of GNNQQNY aggregation. A buffer consisting of 10mM KPi and 300mM NaCl at pH 7.4 with agitation was determined to be ideal for the aggregation. The Gly1 of the peptide was shown to be of importance, potentially aiding in the formation of larger oligomeric structures. Preliminary data shows that Leu2 and Leu6 mutants slow the aggregation rate, possibly due to their inability to form hydrogen bonds. Although the rate was

decreased, Leu2 actually seemed to increase the amount of aggregate formed at thermodynamic equilibrium, which may be caused by the increased hydrophobicity of the peptide. Meanwhile, the Leu3 mutant seems to drastically increase the rate of aggregation. While there are many experiments left to perform on the peptide GNNQQNY, this research has laid an important groundwork for future studies into the aggregation of small peptides.

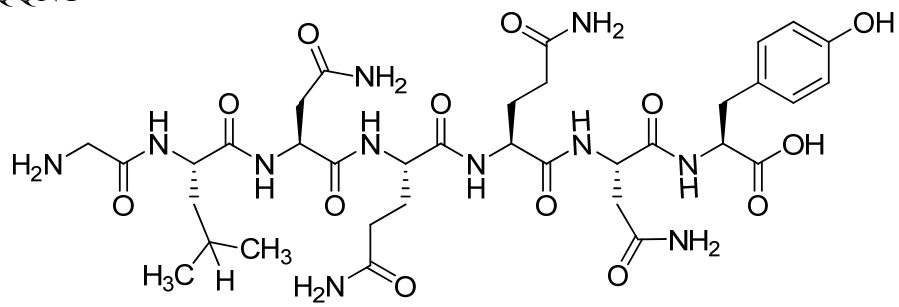
## Figures and Tables:

Figure 1: GNNQQNY and Leucine Mutant Structures.

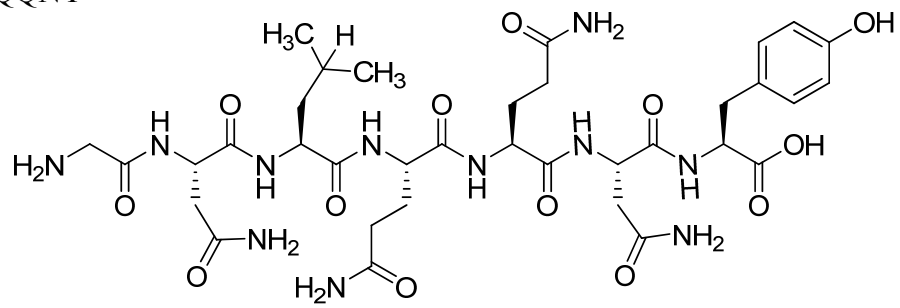
A) GNNQQNY



B) GLNQQNY



C) GNLQQNY



D) GNNQQLY

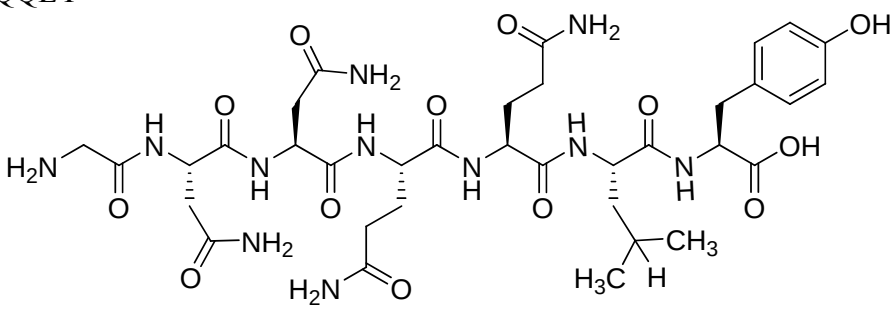
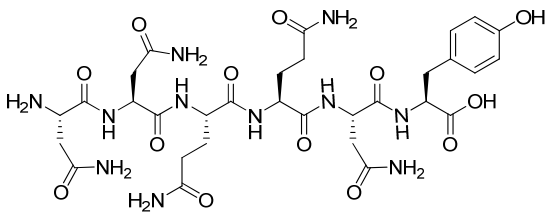
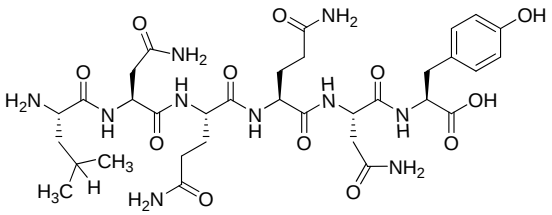


Figure 2: NNQQNY and Mutant Structures.

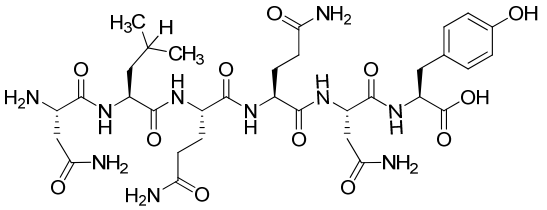
A) NNQQNY



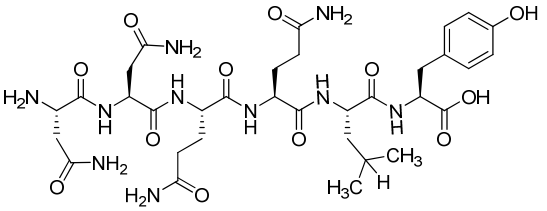
B) LNQQNY



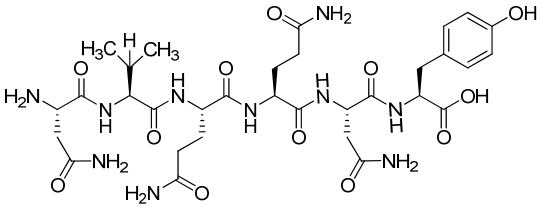
C) NLQQNY



D) NNQQLY



E) NVQQNY



F) NNQQVY

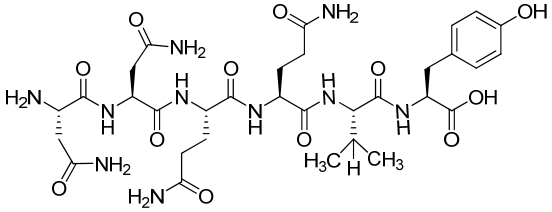


Figure 3: Proton NMR Spectra for GNNQQNY in $D_2O$.

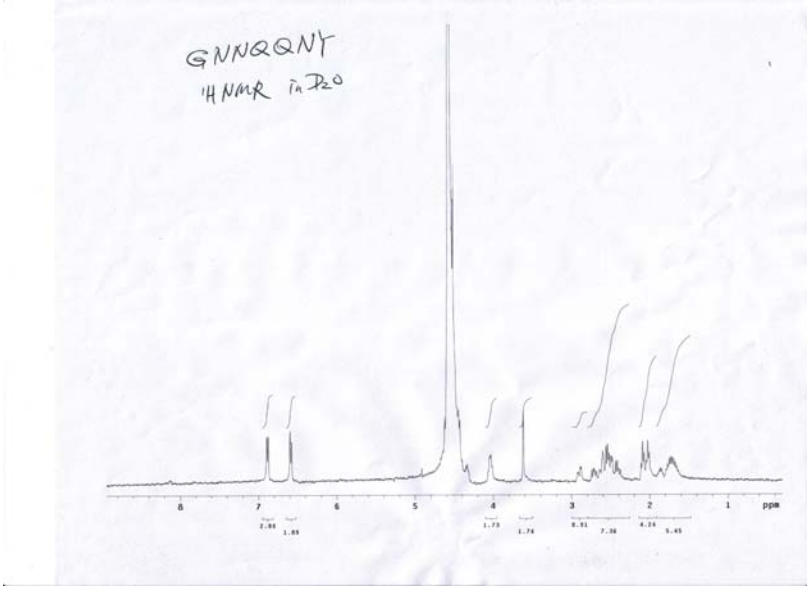
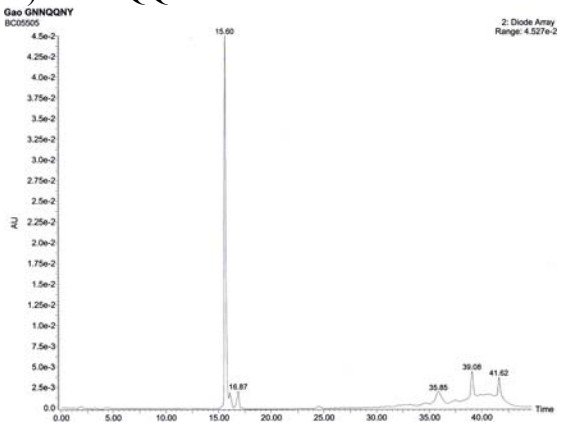
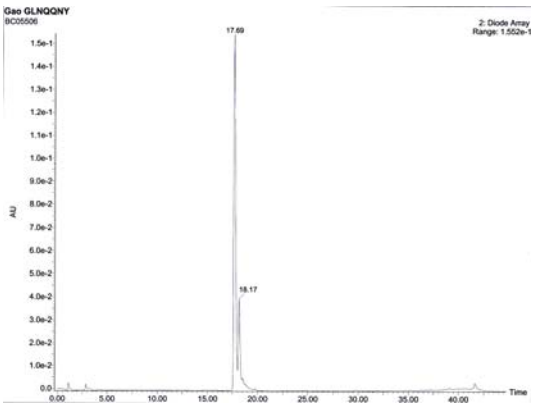



Figure 4: LCMS Data for GNNQQNY and its Mutants.

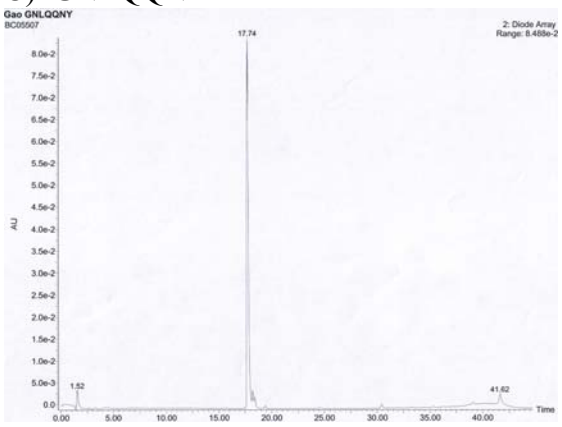
A) GNNQQNY




B) GLNQQNY




C) GNLQQNY




D) GNNQQLY

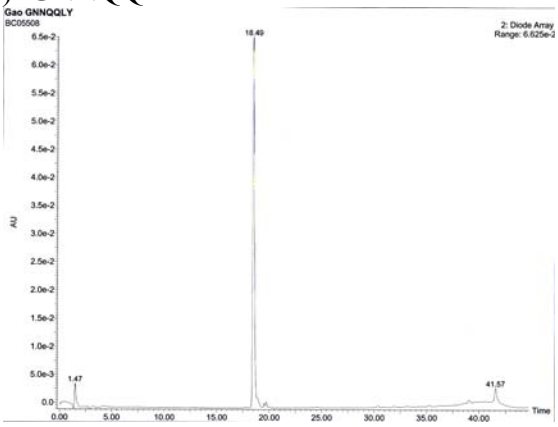
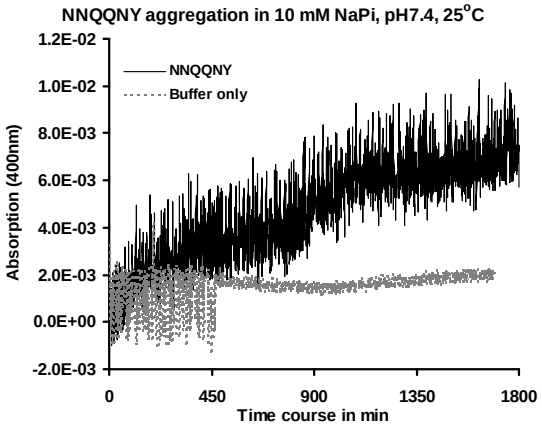
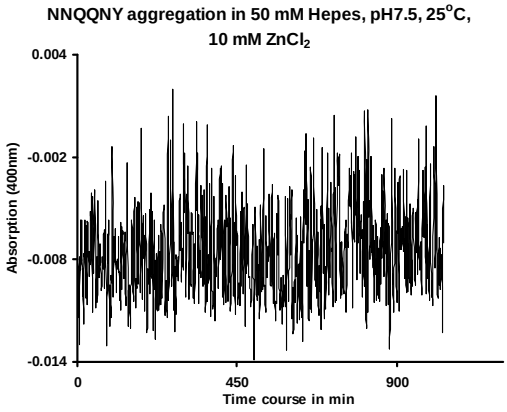


Figure 5: NNQQNY Turbidity Kinetics. *The traces were unclear and difficult to study, showing only small increases in absorption over long periods of time.*

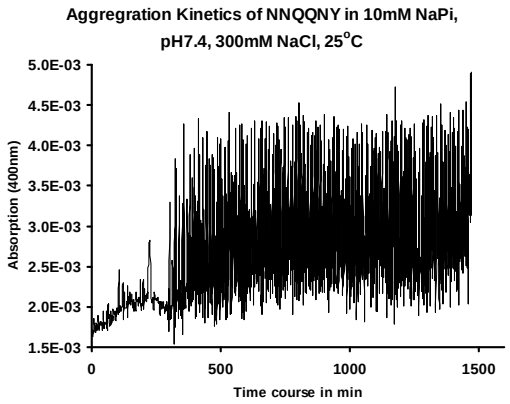
A)




B)




C)




D)

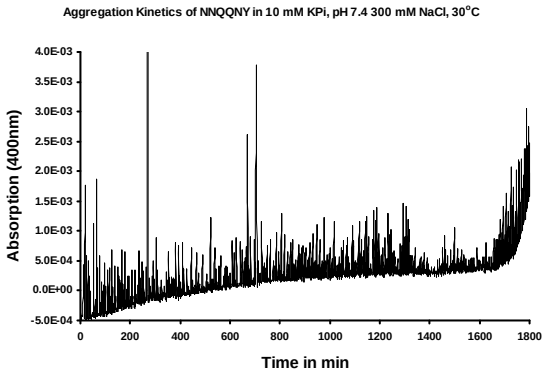
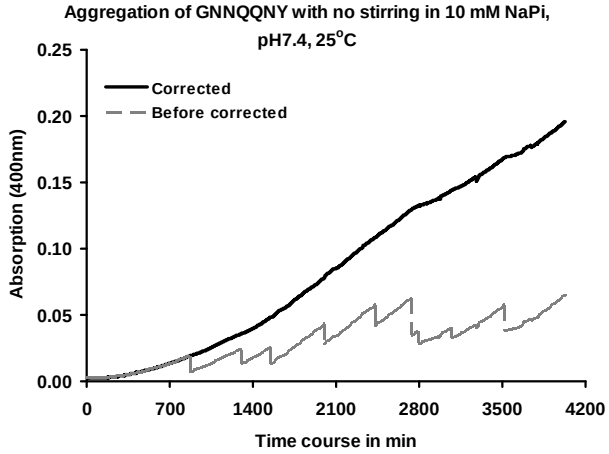
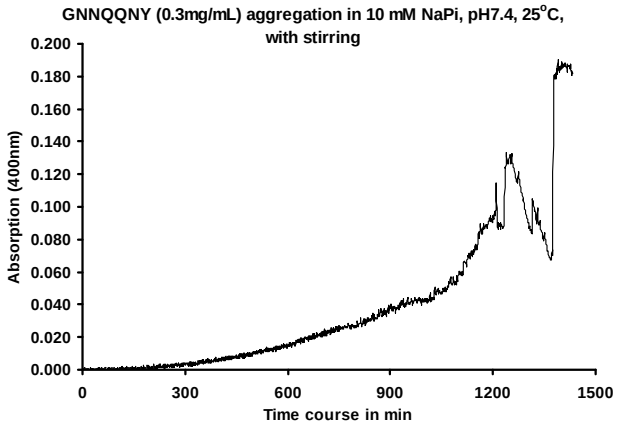


Figure 6: GNNQQNY Turbidity Kinetics. *The jagged areas in the traces are caused by the precipitation of the aggregates, which alters the absorption due to the loss of peptides from solution.*

A)




B)




C)

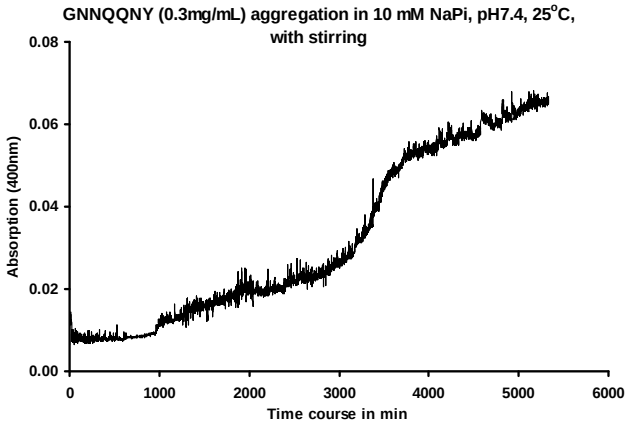
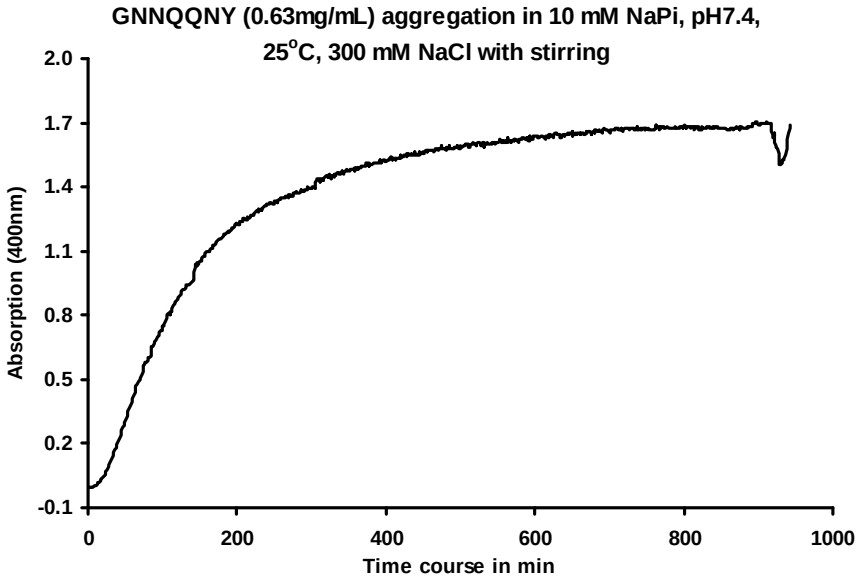


Figure 7: Turbidity of GNNQQNY at Different Concentrations. *The traces in B show a break at 1500min due to the cuvettes being removed from the UV/VIS to be examined visually. This action shakes the solution, suspending any precipitate and increasing absorption.*

A)




B)

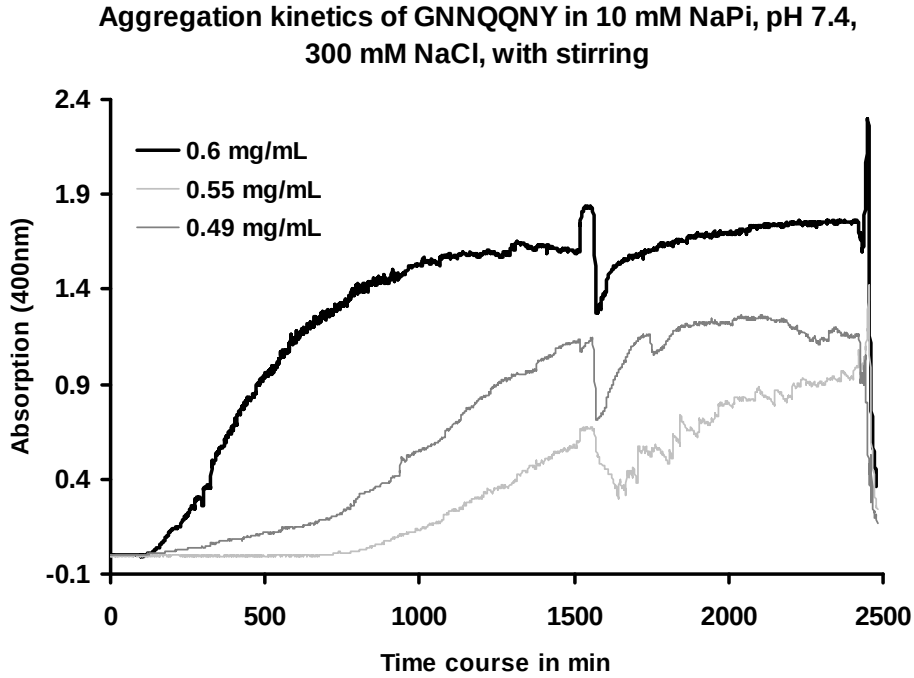
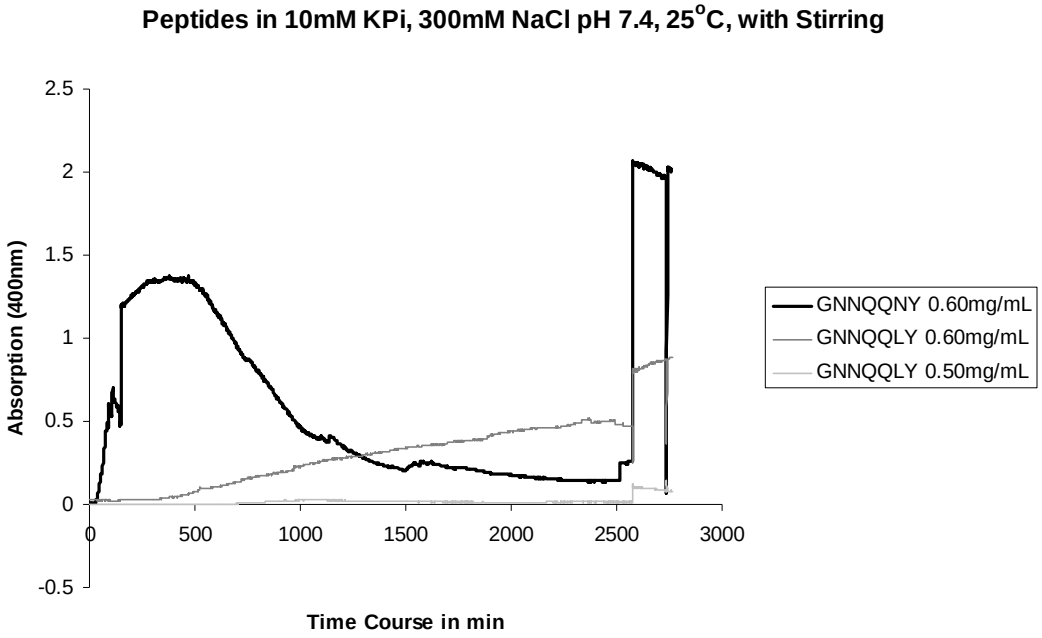


Figure 8: GNNQQNY Mutant Turbidity Kinetics. *The sudden jumps seen in the traces are due to the cuvettes being removed from the UV/VIS to be examined visually. This action shakes the solution, suspending any precipitate and increasing absorption.*

A)




B)

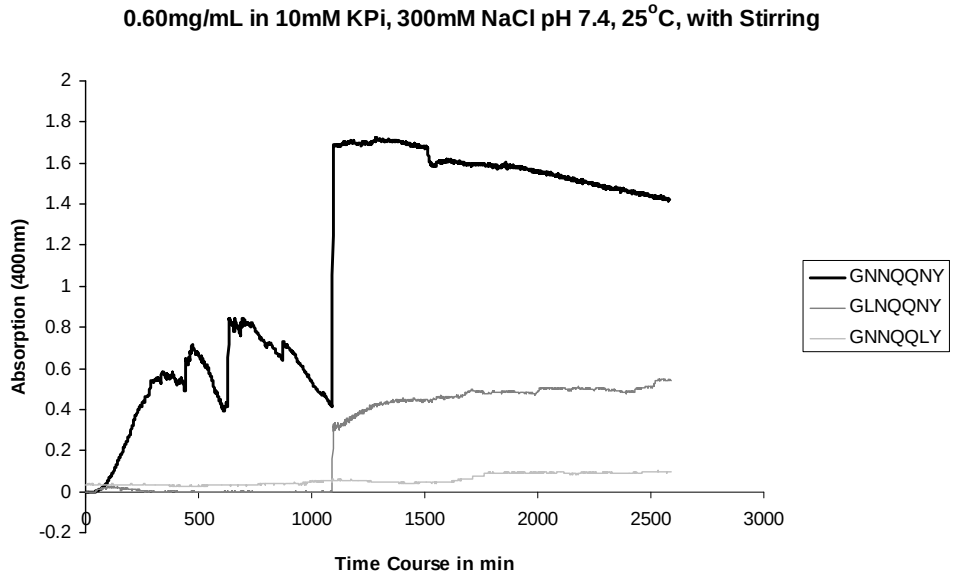


Figure 9: Triplicate GNNQQNY Turbidity Kinetics. *Trial 2 shows a decrease in absorption after 400min due to aggregates precipitating from the solution. The trace for Trial 3 indicates some aggregation, but remains jagged due to the aggregates precipitating.*

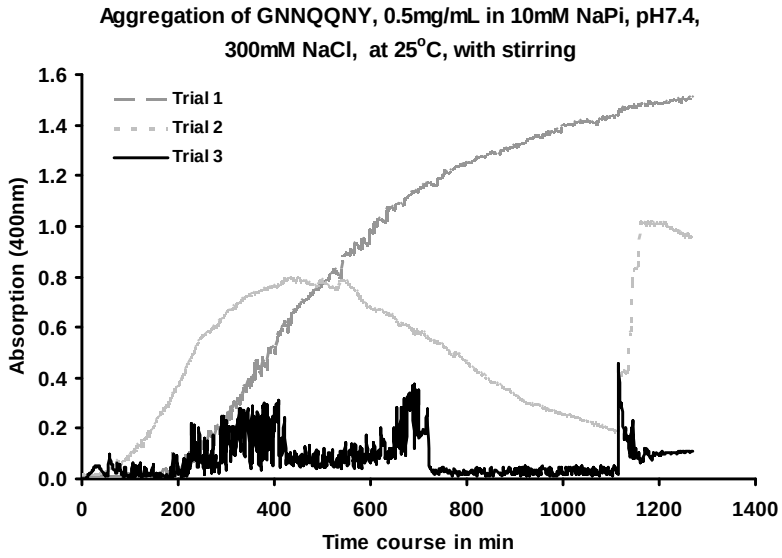



Figure 10: Proton NMR Spectra from Thioflavin T.

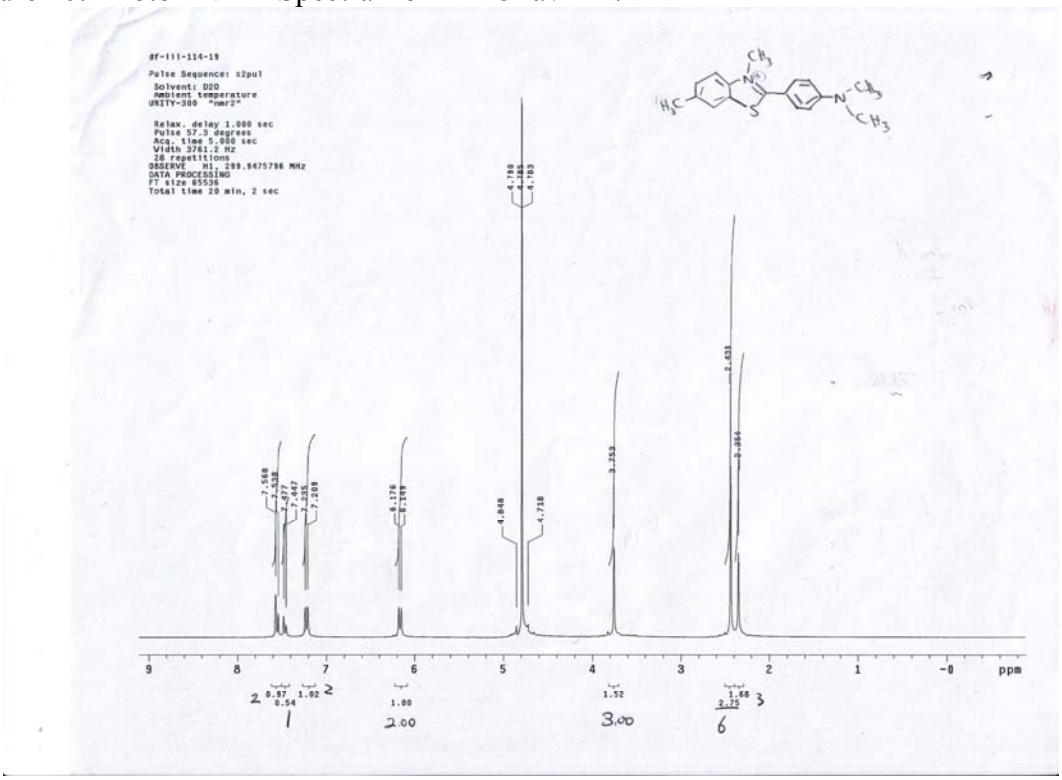
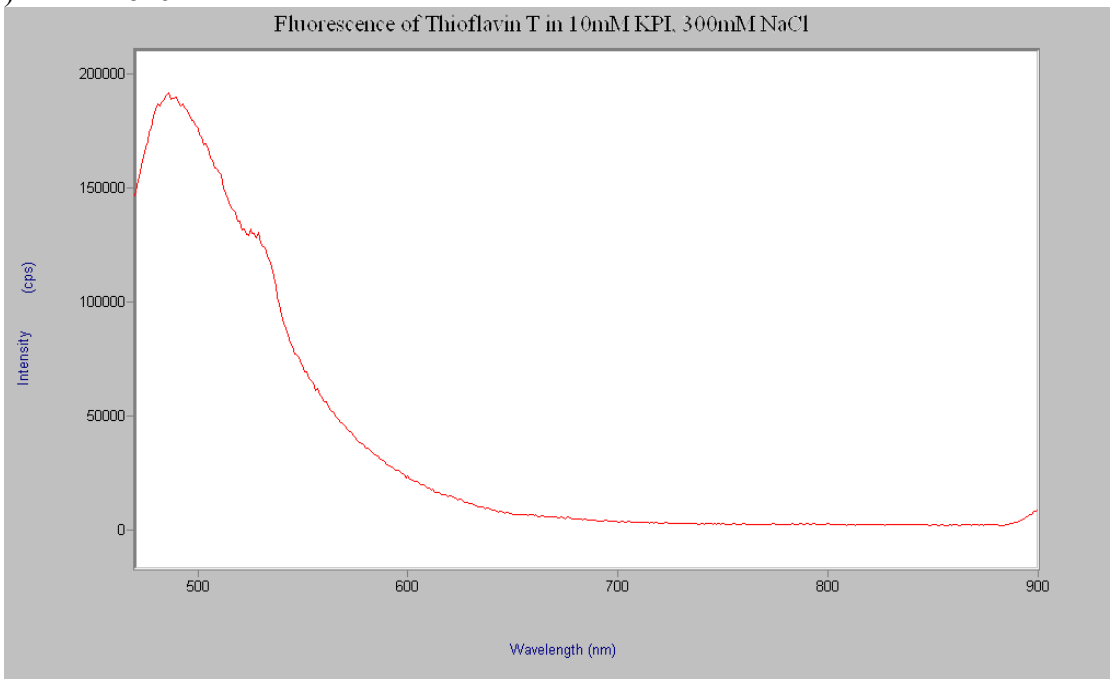


Figure 11: Fluorescent Spectrometry of GNNQQNY Aggregates.

A) ThT Alone




B) 0.63mg/mL GNNQQNY

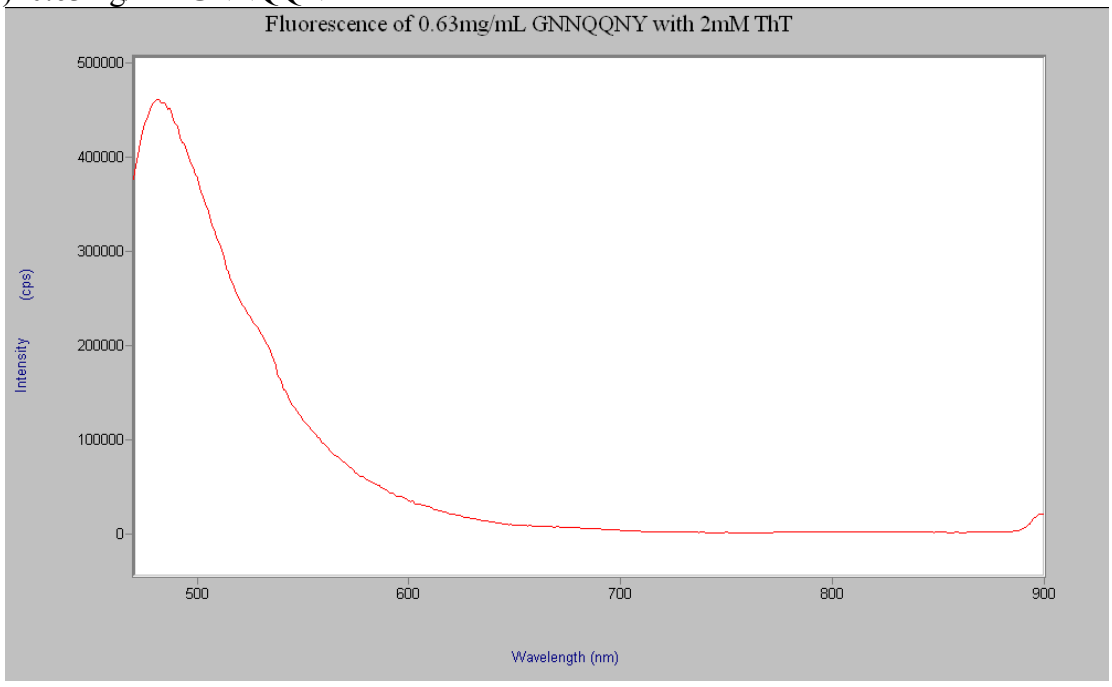
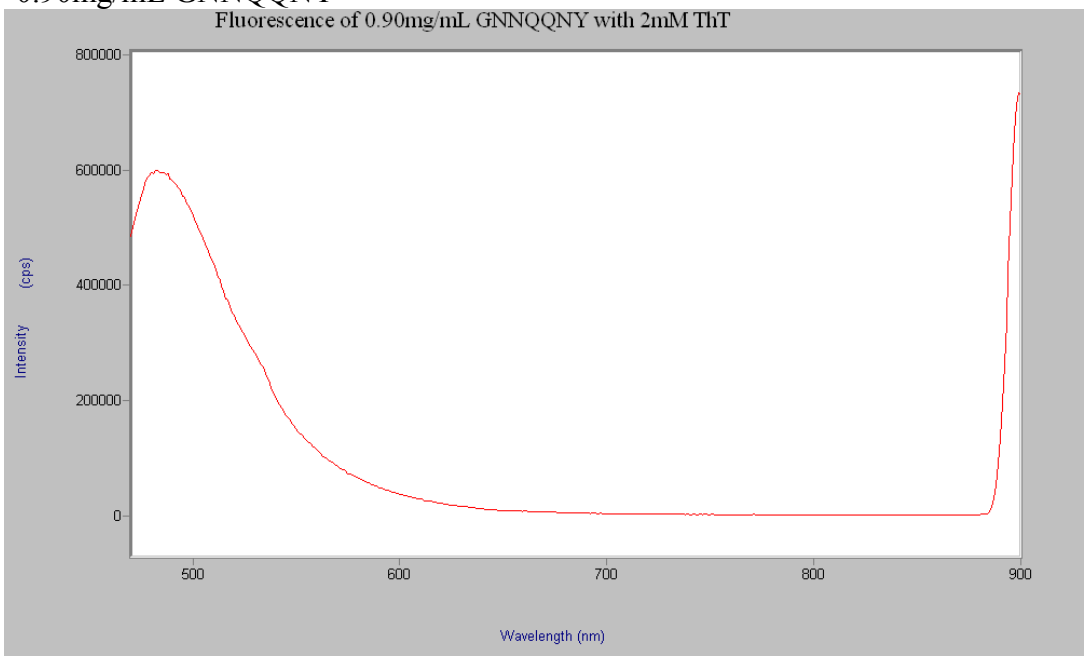


Figure 11 Continued.

C) 0.90mg/mL GNNQQNY




D) 1.50mg/mL GNNQQNY

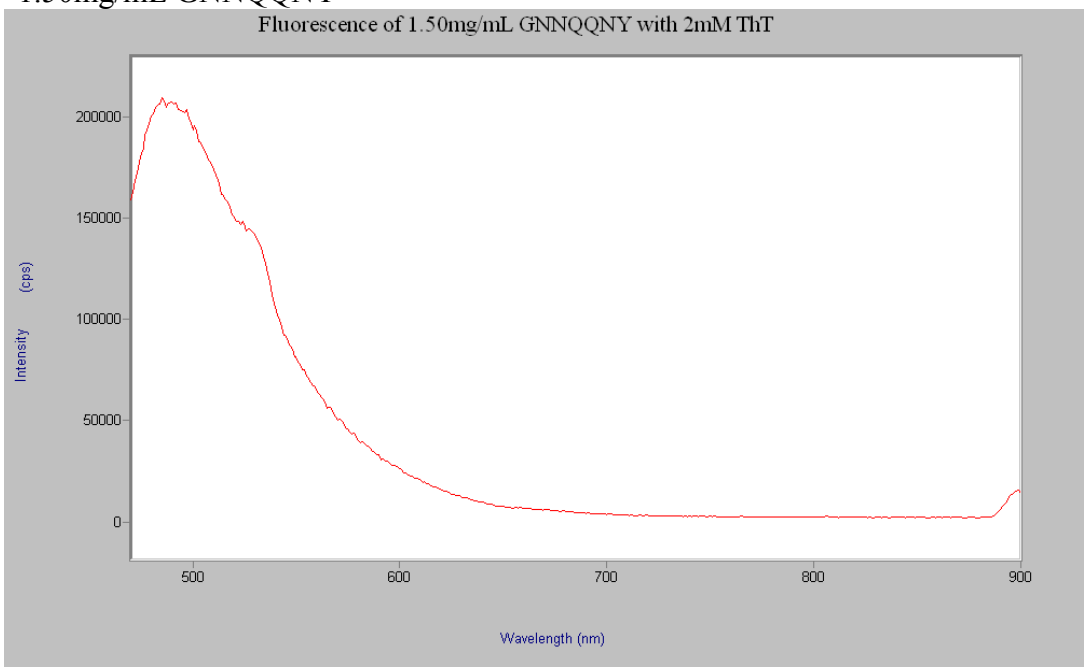
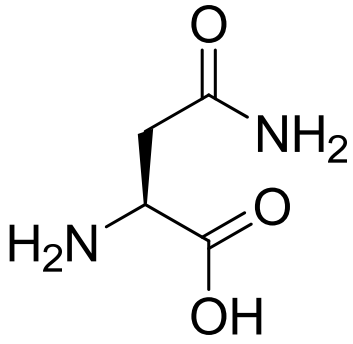
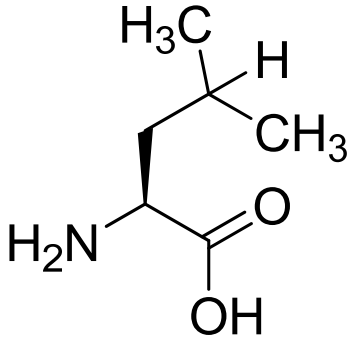


Figure 12: Alternative Amino Acids.

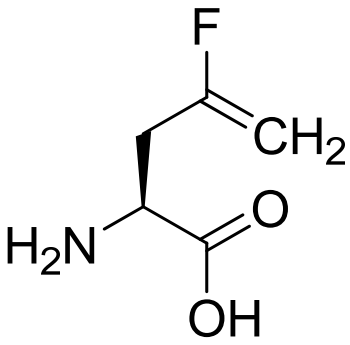
A) Natural Asparagine



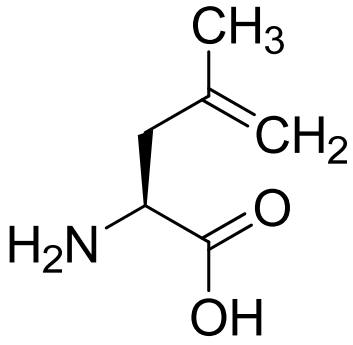

B) Natural Leucine




C)




D)




E)

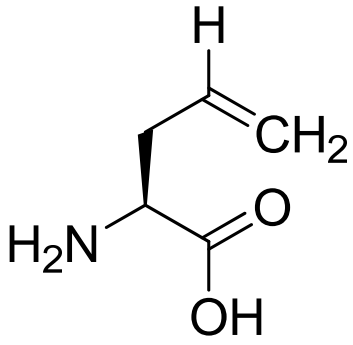


Table 1: Percent of GNNQQNY Aggregated at varied pH

| *Buffer* | *Initial Monomer Concentration (mg/mL)* | *Final Monomer Concentration (mg/mL)* | *Percent of Peptide Aggregated* |
|---|---|---|---|
| 0.01 HCl pH 2 | 0.122 | 0.116 | 4.91% |
| Tris pH 4 | 0.173 | 0.162 | 6.36% |
| Tris pH 6 | 0.156 | 0.130 | 16.7% |
| Tris pH 8 | 0.160 | 0.103 | 35.6% |
| Tris pH 10 | 0.157 | 0.131 | 16.6% |
| 300mM NaCl pH 7.4 | 0.134 | 0.081 | 39.6% |

Table 2: Percent of GNNQQNY Aggregated at varied pH,
with Initial Peptide Concentration of 1mM

| *Buffer* | *Monomer Concentration without agitation (mM)* | *Monomer Concentration after agitation (mM)* | *Percent of Peptide Aggregated* |
|---|---|---|---|
| 0.01 HCl pH2 | 1.05 | *no reading* | 0% |
| Tris pH 4 | 0.980 | 0.432 | 56.8% |
| Tris pH 6 | 0.976 | 1.09 | 0% |
| Tris pH 8 | 0.996 | 0.992 | 0.8% |
| NaPi pH 8 | 1.01 | 1.04 | 0% |
| Tris pH 10 | 0.957 | 0.970 | 3.0% |
| 300mM NaCl pH 7.4 | 0.905 | 0.919 | 8.1% |

*Note: The agitation times for each sample were inconsistent. Also, no measurement was taken for the sample at pH 2 after agitation, due to possible contamination of the sample.*

Table 3: Peptide Monomer Concentrations
in 10mM KPi, 300mMNaCl Buffer

| *Experiment* | *Peptide* | *Concentration Before Aggregation (mg/mL)* | *Concentration After Aggregation (mg/mL)* | *Percent of Peptide Aggregated* |
|---|---|---|---|---|
| Concentration Optimization | GNNQQNY | 0.63 | 0.28 | 56% |
| | GNNQQNY | 0.90 | 0.38 | 58% |
| | GNNQQNY | 1.50 | 0.44 | 71% |
| First Mutant Experiment | GNNQQNY | 0.60 | 0.24 | 60% |
| | GNNQQLY | 0.60 | 0.26 | 57% |
| | GNNQQLY | 0.50 | 0.41 | 18% |
| Second Mutant Experiment | GNNQQNY | 0.60 | 0.26 | 57% |
| | GLNQQNY | 0.60 | 0.14 | 77% |
| | GNNQQLY | 0.60 | 0.40 | 33% |
| Triplicate Experiment | GNNQQNY | 0.50 | 0.25 | 50% |
| | GNNQQNY | 0.50 | 0.26 | 48% |
| | GNNQQNY | 0.50 | 0.24 | 52% |

*Note: The Post-Aggregation Monomer Concentrations were measured several weeks after the initial aggregation experiment. Since the kinetic data did not show the end-plateau for every mutant, it is difficult to say when the reactions reached completion, or if they were complete when the measurements were taken.*

# Appendix:

**Amino Acid Abbreviations**

| *Amino Acid* | *Three Letter* | *One Letter* |
|---|---|---|
| Alanine | Ala | A |
| Asparagine | Asn | N |
| Aspartate | Asp | D |
| Arginine | Arg | R |
| Cysteine | Cys | C |
| Glutamate | Glu | E |
| Glutamine | Gln | Q |
| Glycine | Gly | G |
| Histidine | His | H |
| Isoleucine | Ile | I |
| Leucine | Leu | L |
| Lysine | Lys | K |
| Methionine | Met | M |
| Phenylalane | Phe | F |
| Proline | Pro | P |
| Serine | Ser | S |
| Threonine | Thr | T |
| Tryptophan | Trp | W |
| Tyrosine | Tyr | Y |
| Valine | Val | V |

## Acknowledgments:

I would especially like to thank Dr. Jianmin Gao for taking me into his lab and giving me my own project. I was only able to complete this research due to his constant support and guidance.

I would also like to thank Dr. Tao Ye for helping so much with the aggregation experiments. I could not have drawn any conclusions from this study without the time and effort he put in.

I want to thank Hong Zheng for his help with the peptide syntheses. I would also like to thank the rest of the Gao Group for all of their help, ranging from experimental ideas to simply finding the beakers.

Thanks are also due to my roommates and friends who put up with me during the writing process and who were always around to entertain me when I needed a break. In particular I would like to thank Tom Keller, who helped to proofread the final product.

Finally, I would like to thank my family and Diane, who have always supported me in all of my efforts.

Thanks everyone, I could not have done any of this without you.

*Kevin Lebo, April 30$^{th}$ 2008*

## References:


1. Fink, A.L. (1998) Protein Aggregation: folding aggregates, inclusion bodies and amyloid. *Folding and Design.* 3:R9-R23

2. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A., and Ventura, S. (2005) Amyloid-like Properties of Bacterial Inclusion Bodies. *Journal of Molecular Biology.* 347:1025-1037

3. Sipe, J.D. and Cohen, A.S. (2000) Review: History of the Amyloid Fibril. *Journal of Structural Biology.* 130:88-98

4. Sunde, M., Serpell, L.C., Bartlam, M., Fraser, P.E., Pepys, M.B., and Blake, C.F. (1997) Common Core Structure of Amyloid Fibrils by Synchrotron X-ray Diffraction. *Journal of Molecular Biology.* 273:729-739

5. Balbirnie, M., Grothe, R., and Eisenberg, D.S. (2001) An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated β-sheet structure for amyloid. *Proceedings of the National Academy of Sciences.* 98:2375-2380

6. Benseny-Cases, N., Cocera, M., and Cladera, J. (2007) Conversion of non-fibrillar β-sheet oligomers into amyloid fibrils in Alzheimer's disease amyloid peptide aggregation. *Biochemical and Biophysical Research Communications*. 361:916-921

7. Konno, T. (2001) Amyloid-Induced Aggregation and Precipitation of Soluble Proteins: An Electrostatic Contribution of the Alzheimer's β(25-35) Amyloid Fibril. *Biochemistry.* 40:2148-2154

8. Perczel, A., Hudaky, P., and Palfi, V.K. (2007) Dead-End Street of Protein Folding: Thermodynamic Rationale of Amyloid Fibril Formation. *Journal of the American Chemical Society.* 129:14959-14965

9. Krzewska, J. and Melki, R. (2006) Molecular chaperones and the assembly of the prion Sup35p, an *in vitro* study. *The European Molecular Biology Organization Journal.* 25:822-833

10. Zhang, Z., Chen, H., Bai, H., and Lai, L. (2007) Molecular Dynamics Simulations on the Oligomer-Formation Process of the GNNQQNY Peptide from Yeast Prion Protein Sup35. *Biophysical Journal*. 93:1484-1492

11. Michelitsch, M.D. and Weissman, J.S. (2000) A census of the glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel proteins. *Proceedings of the National Academy of Sciences.* 97:11910-11915

12. Kowalewski, T. and Holtzman, D.M. (1999) *In Situ* atomic force microscopy study of Alzheimer's β-amyloid peptide on different substrates: New insights into mechanism of β-sheet formation. *Proceedings of the National Academy of Sciences. USA.* 96:3688-3693

13. Bucciantini, M., Calloni, G., Chiti, F., Formigli, L., Nosi, D., Dobson, C.M., and Stefani, M. (2004) Prefibrillar Amyloid Protein Aggregates Share Common Features of Cytotoxicity. *The Journal of Biological Chemistry.* 279:31374-31382

14. Thundimadathil, J., Roeske, R.W., Jiang, H., and Guo, L. (2005). Aggregation and Porin-like Channel Activity of a β-Sheet Peptide. *Biochemistry.* 44:10259-10270

15. Gonzalez, M.R., Bischofberger, M., Pernot, L., van der Goot, F.G., and Freche, B. (2008). Review: Bacterial pore-forming toxins: The whole story? *Cellular and Molecular Life Sciences*. 65:493-507

16. Soto, P., Baumketner, A., and Shea, J. (2006) Aggregation of polyalanine in a hydrophobic environment. *The Journal of Chemical Physics.* 124:134904

17. Matsuzaki, K. (2007) Physiochemical interaction of amyloid β-peptide with lipid bilayers. *Biochimica et Biophysica Acta.* 1768:1935-1942

18. Reches, M., Porat, Y., and Gazit, E. (2002) Amyloid Fibril Formation by Pentapeptide and Tetrapeptide Fragments of Human Calcitonin. *The Journal of Biological Chemistry.* 277:35475-354800

19. McLaurin, J., Yang, D.S., Yip, C.M., and Fraser, P.E. (2000) Review: Modulating Factors in Amyloid- β Fibril Formation. *Journal of Structural Biology*. 130:259-270

20. Dodart, J., Bales, K.R., Gannon, K.S., Greene, S.J., DeMattos, R.B., Mathis, C., DeLong, C.A., Wu, S., Wu, X., Holtzman, D.M., and Paul, S.M. (2002) Immunization reverses memory deficits without reducing brain Aβ burden in Alzheimer's disease model. *Nature Neuroscience*. 5:452-457

21. Findeis, M.A., Musso, G.M., Arico-Muendel, C.C., Benjamin, H.W., Hundal, A.M., Lee, J., Chin, J, Kelley, M., Wakefield, J, Hayward, N.J., and Molineaux, S.M. (1999) Modified-Peptide Inhibitors of Amyloid β-Peptide Polymerization. *Biochemistry*. 38:6791-6800

22. Etienne, M.A., Aucoin, J.P., Fu, Y., McCarley, R.L., and Hammer, R.P. (2006) Stoichiometric Inhibition of Amyloid β-Protein Aggregation with Peptides Containing Alternating α,α-Disubstituted Amino Acids. *Journal of the American Chemical Society.* 128:3522-3523

23. Monsellier, E. and Chiti, F. (2007) Prevention of amyloid-like aggregation as a driving force of protein evolution. *European Molecular Biology Organization Reports.* 8:737-742

24. Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C.M. (2003) Rationalization of the effects of mutations on peptide aggregation rates. *Letters to Nature*. 424:805-808

25. Lopez de la Paz, M., and Serrano, L. (2004) Sequence determinants of amyloid fibril formation. *Proceedings of the National Academy of Sciences.* 101:87-92

26. Street, A.G., and Mayo, S.L. (1999) Intrinsic β-sheet propensities result from van der Waals interactions between side chains and the local backbone. *Proceedings of the National Academy of Sciences.* 96:9074-9076

27. Wurth, C., Guimard, N.K., and Hect, M.H. (2002) Mutations that Reduce Aggregation of the Alzheimer's Aβ42 Peptide: an Unbiased Search for the Sequence Determinants of Aβ Amyloidogenesis. *Journal of Molecular Biology.* 319:1279-1290

28. Wetzel, R. (2006) Kinetics and Thermodynamics of Amyloid Fibril Assembly. *Accounts of Chemical Research.* 39:671-679

29. Kim, W. and Hecht, M.H. (2006) Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Aβ42 peptide. *Proceedings of the National Academy of Sciences.* 43:15824-15829

30. Sanchez de Groot, N., Aviles, F.X., Vendrell, J, and Ventura, S. (2006) Mutagenesis of the central hydrophobic cluster in Aβ42 Alzheimer's peptide. *The FEBS Journal.* 273:658-668

31. Wu, C., Lei, H., and Duan, Y. (2005) Elongation of Ordered Peptide Aggregate of an Amyloidogenic Hexapeptide NFGAIL Observed in Molecular Dynamics Simulation with Explicit Solvent. *Journal of the American Chemical Society*. 127:13530-13537

32. Gsponer, J., Haberthur, U., and Caflisch, A. (2003) The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proceedings of the National Academy of Sciences*. 100:5154-5159

33. Lipfert, J., Franklin, J., Wu, F., and Doniach, S. (2005) Protein Misfolding and Amyloid Formation for the Peptide GNNQQNY from Yeast Prion Protein Sup35: Simulations by Reaction Path Annealing. *Journal of Molecular Biology.* 349:648-658

34. Nelson, R, Sawaya, M.R., Balbirnie, M., Madsen, A., Riekel, C., Grothe, R., and Eisenberg, D. (2005) Structure of the cross-β spine of amyloid-like fibrils. *Nature*. 435:773-778